

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 09:18:16 ; Search time 1242.59 seconds  
(without alignments)  
13865.275 Million cell updates/sec

Title: US-09-905-114-1  
Perfect score: 592  
Sequence: 1 gagaagcctaagcgaattc.....gtctgacccaagccctgg 592

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GeneBml: \*  
1: gb\_ba: \*  
2: gb\_hcg: \*  
3: gb\_in: \*  
4: gb\_ov: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vl: \*  
15: em\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_mu: \*  
20: em\_or: \*  
21: em\_ov: \*  
22: em\_ov: \*  
23: em\_pat: \*  
24: em\_ph: \*  
25: em\_pl: \*  
26: em\_ro: \*  
27: em\_sts: \*  
28: em\_un: \*  
29: em\_vl: \*  
30: em\_hcg\_hum: \*  
31: em\_hcg\_inv: \*  
32: em\_hcg\_other: \*  
33: em\_hcg\_mus: \*  
34: em\_hcg\_pin: \*  
35: em\_hcg\_rod: \*  
36: em\_hcg\_mam: \*  
37: em\_hcg\_vrt: \*  
38: em\_sy: \*  
39: em\_hugo\_hum: \*  
40: em\_hugo\_mus: \*  
41: em\_hugo\_other: \*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	476.8	80.5	1023	AX329742	AX329742 Sequence
2	476.8	80.5	1023	AX332286	AX332286 Sequence
3	476.8	80.5	1023	AX332789	AX332789 Sequence
4	476.8	80.5	1023	AX410746	AX410746 Sequence
5	476.8	80.5	1023	HSU56814	US6814 Human DNase
6	476.8	80.5	1067	BC015831	BC015831 Homo sapi
7	476.8	80.5	1079	AF047354	AF047354 Homo sapi
8	475.2	80.3	1108	HSU75744	U75744 Homo sapien
9	459.2	77.6	1131	AF039852	AF039852 Rattus no
10	459.2	77.6	1417	RNU75689	U75689 Rattus norv
11	452.8	76.5	1208	AR047845	AR047845 Sequence
12	452.8	76.5	1208	E11687	E11687 cDNA encodi
13	435.2	73.5	941	MMU76110	U76110 Mus musculu
14	435.2	73.5	1124	AF047355	AF047355 Mus muscu
15	433.6	73.2	2110	BC012671	BC012671 Mus muscu
16	387.2	65.4	1223	AX400016	AX400016 Sequence
17	280.8	29.3	1295	AF059612	AF059612 Xenopus 1
18	173.2	29.3	1295	BTNPANDAI	AJ001538 Bos tauru
19	172.2	29.1	868	AF311922	AF311922 Bos tauru
20	171.4	29.0	392	AX330074	AX330074 Sequence
21	171.4	29.0	392	AX407983	AX407983 Sequence
22	163.2	27.6	1210	AB038776	AB038776 Rana cete
23	160.6	27.1	1283	AB041732	AB041732 Cynops py
24	154	26.0	1161	D82875	D82875 Oryctolagus
25	152.2	25.7	1312	BC014718	BC014718 Mus muscu
26	152.2	25.7	1337	BC030394	BC030394 Mus muscu
27	151	25.5	1297	AB048832	AB048832 Sus scrof
28	150.8	25.5	1156	RNDES0X1	X56060 Rat mRNA fo
29	150.8	25.5	1184	AF397151	AF397151 Rattus no
30	150.8	25.5	1313	AF397150	AF397150 Rattus no
31	150.6	25.4	1180	D83038	D83038 Mus musculu
32	149.4	25.2	1039	AX3030575	AX3030575 Sequence
33	149.4	25.2	1039	AX030591	AX030591 Sequence
34	149.4	25.2	1039	AX268671	AX268671 Sequence
35	149.4	25.2	1039	HMUNDAISEI	M55983 Human DNase
36	149.4	25.2	1149	BC029437	BC029437 Homo sapi
37	149.4	25.2	2054	HS2898844	AJ298844 Homo sapi
38	148.4	25.1	1066	AB013751	AB013751 Gallus ga
39	147.8	25.0	783	AX268673	AX268673 Sequence
40	147.8	25.0	858	AX268675	AX268675 Sequence
41	147.8	25.0	1039	AR019468	AR019468 Sequence
42	147.8	25.0	1548	AX268721	AX268721 Sequence
43	147.8	25.0	1548	AX268722	AX268722 Sequence
44	147.8	25.0	1554	AX268701	AX268701 Sequence
45	147.8	25.0	1554	AX268702	AX268702 Sequence

#### ALIGNMENTS

RESULT 1  
AX329742 LOCUS AX329742 1023 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 251 from Patent WO0194629.  
ACCESSION AX329742  
VERSION AX329742.1 GI:18102720  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
1 Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endreess, G.,  
Horrigan, S., Soppet, D.R. and Weaver, Z.  
Cancer gene determination and therapeutic screening using signature

Pred. No. is the number of results predicted by chance to have a

gene sets  
Patent: WO 0194629-A 251 13-DEC-2001;  
Avalon Pharmaceuticals (US)  
Location/Qualifiers  
/organism="Homo sapiens"  
1..1023  
/db\_xref="taxon:9606"  
BASE COUNT 312 a 244 c 240 g 227 t  
ORIGIN

Query Match 80.5%; Score 476.8; DB 6; Length 1023;  
Best Local Similarity 87.8%; Pred. No. 1.7e-119;  
Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 1 GAGAGCTAAACCGAAATTCAGAAAGGCATACATCAACTATGTGATGTCTCGC 60  
DB 241 GAGAGCTGAACAGAAATTCAGAGGAGGCATACATCAACTATGTGATGTCTCGC 300  
QY 61 CTGGAAGAAACACATATATAAGAACAGTATGCTTCTCTATATAAGAAAGCTAGTGCT 120  
DB 301 CTGGAAGAAACACATATATAAGAACATATGCTTCTCTCAAGAAAGCTGTGTCT 360  
QY 121 GTAAACAAAGCTACCTCTACCAAGCTATCAGGCTGGAGAGCGAGATGTGTTTCCAGG 180  
DB 361 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGAGAGCGAGATGTGTTTCCAGG 420  
QY 181 GAACCTTTGTGCTGTGTTCCAGTCAACCTACACCGCTGTCAAGGACTTCGTGATGTC 240  
DB 421 GAGCCCTTTGTGCTGTGTTCCATCTCCCCACACTGCTGTCAAGACTTCGTGATGTC 480  
QY 241 CCCCTGCACACACCCCTTGAGACATCCGTTAGAGATGTGATGCTGGCTGATGCTAC 300  
DB 481 CCCCTGCACACACCCCTTGAGACATCCGTTAGAGATGTGATGCTGGCTGATGCTAC 540  
QY 301 ACAGATGTGAACCGTGTGGAATGCAGAGAAATTCATTTTCATGGTGACTTCAATGCT 360  
DB 541 ACGGACGTGAACACCGCTGGAAGCGGAGAAATTCATTTTCATGGTGACTTCAATGCT 600  
QY 361 GGCTGCAGCTACGTCCTCCCAAGAGCGCTGGAAGGACATCCGCTGAGAGCGAGCCCAAG 420  
DB 601 GGCTGCAGCTACGTCCTCCCAAGAGCGCTGGAAGGACATCCGCTGAGAGCGAGCCCAAG 660  
QY 421 TTCGTTGGCTGATCGGGGACCAAGAGACACACCGCTCAAGAGAGACACAACTCGGCC 480  
DB 661 TTTGTTGGCTGATCGGGGACCAAGAGACACACCGCTCAAGAGAGACACAACTCGGCC 720  
QY 481 TATGACAGGATCGCTTAGAGACAAATATTTGTCATCTGCTGCTCAATCAAAAC 540  
DB 721 TATGACAGGATGTGCTTAGAGACAAATATCGTCAGTTCTGTTGTTCCCAAGTCAAAAC 780  
QY 541 CTCGCTTTGATTTCCAGAAAGCTTACAGGTTGCTGAATCGAAGGCCCTGG 592  
DB 781 AGTGTGTTTGACTTCCAGAAAGCTTACAGCTGACTGAAGAGGAGGCCCTGG 832

RESULT 2  
AX332286  
LOCUS AX332286 1023 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 2795 from Patent WO0194629.  
ACCESSION AX332286  
VERSION AX332286.1 GI:18122920  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE  
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,  
Horrihan, S., Soppet, D.R. and Weaver, Z.  
TITLE Cancer gene determination and therapeutic screening using signature  
gene sets  
JOURNAL Patent: WO 0194629-A 2795 13-DEC-2001;  
Avalon Pharmaceuticals (US)  
FEATURES  
source

Location/Qualifiers  
1..1023  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 312 a 244 c 240 g 227 t  
ORIGIN

Query Match 80.5%; Score 476.8; DB 6; Length 1023;  
Best Local Similarity 87.8%; Pred. No. 1.7e-119;  
Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 1 GAGAGCTAAACCGAAATTCAGAAAGGCATACATCAACTATGTGATGTCTCGC 60  
DB 241 GAGAGCTGAACAGAAATTCAGAGGAGGCATACATCAACTATGTGATGTCTCGC 300  
QY 61 CTGGAAGAAACACATATAAGAACAGTATGCTTCTCTATATAAGAAAGCTAGTGCT 120  
DB 301 CTGGAAGAAACACATATAAGAACATATGCTTCTCTCAAGAAAGCTGTGTCT 360  
QY 121 GTAAACAAAGCTACCTCTACCAAGCTATCAGGCTGGAGAGCGAGATGTGTTTCCAGG 180  
DB 361 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGAGAGCGAGATGTGTTTCCAGG 420  
QY 181 GAACCTTTGTGCTGTGTTCCAGTCAACCTACACCGCTGTCAAGGACTTCGTGATGTC 240  
DB 421 GAGCCCTTTGTGCTGTGTTCCATCTCCCCACACTGCTGTCAAGACTTCGTGATGTC 480  
QY 241 CCCCTGCACACACCCCTTGAGACATCCGTTAGAGATGTGATGCTGGCTGATGCTAC 300  
DB 481 CCCCTGCACACACCCCTTGAGACATCCGTTAGAGATGTGATGCTGGCTGATGCTAC 540  
QY 301 ACAGATGTGAACCGTGTGGAATGCAGAGAAATTCATTTTCATGGTGACTTCAATGCT 360  
DB 541 ACGGACGTGAACACCGCTGGAAGCGGAGAAATTCATTTTCATGGTGACTTCAATGCT 600  
QY 361 GGCTGCAGCTACGTCCTCCCAAGAGCGCTGGAAGGACATCCGCTGAGAGCGAGCCCAAG 420  
DB 601 GGCTGCAGCTACGTCCTCCCAAGAGCGCTGGAAGGACATCCGCTGAGAGCGAGCCCAAG 660  
QY 421 TTCGTTGGCTGATCGGGGACCAAGAGACACACCGCTCAAGAGAGACACAACTCGGCC 480  
DB 661 TTTGTTGGCTGATCGGGGACCAAGAGACACACCGCTCAAGAGAGACACAACTCGGCC 720  
QY 481 TATGACAGGATCGCTTAGAGACAAATATTTGTCATCTGCTGCTCAATCAAAAC 540  
DB 721 TATGACAGGATGTGCTTAGAGACAAATATCGTCAGTTCTGTTGTTCCCAAGTCAAAAC 780  
QY 541 CTCGCTTTGATTTCCAGAAAGCTTACAGGTTGCTGAATCGAAGGCCCTGG 592  
DB 781 AGTGTGTTTGACTTCCAGAAAGCTTACAGCTGACTGAAGAGGAGGCCCTGG 832

RESULT 3  
AX332789  
LOCUS AX332789 1023 bp DNA linear PAT 10-JAN-2002  
DEFINITION Sequence 3298 from Patent WO0194629.  
ACCESSION AX332789  
VERSION AX332789.1 GI:18123423  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE  
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,  
Horrihan, S., Soppet, D.R. and Weaver, Z.  
TITLE Cancer gene determination and therapeutic screening using signature  
gene sets  
JOURNAL Patent: WO 0194629-A 3298 13-DEC-2001;  
Avalon Pharmaceuticals (US)  
FEATURES  
source

BASE COUNT 312 a 244 c 240 g 227 t  
ORIGIN

Query Match 80.5%; Score 476.8; DB 6; Length 1023;  
Best Local Similarity 87.8%; Pred. No. 1.7e-119;  
Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

1 GAAACCTAAACGGAATTCAGAAAGGCAATACATACATATGATTTAGCTCCG 60  
DB 241 GAAACCTAAACGGAATTCAGAAAGGCAATACATACATATGATTTAGCTCCG 300  
QY 61 CTGGAAGAAACATATATAAGAAACATATGCTTTCTCTATAAGAAAGCTAGTGTCT 120  
DB 301 CTGGAAGAAACATATATAAGAAACATATGCTTTCTCTATAAGAAAGCTAGTGTCT 360  
QY 121 GTAAACAAAGCTACTCTACACAGACTATCAGGCTGAGAGCCAGATGTTTTCCAGG 180  
DB 361 GTGAAGAGGATTAATCACTACATGACTATCAGGATGAGACCCAGATGTTTTCCAGG 420  
QY 181 GAACCTTTGTGTCTGTTCCAGTACACCTACACCGCTGCAAGGCTCGTGAATGTC 240  
DB 421 GAACCTTTGTGTCTGTTCCAGTACACCTACACCTGCTGTCAAGACTTCTGATTTATC 480  
QY 241 CCCCTGACACACACCCCTGAGACATCCGTTAGAGATGATGAGCTGCTGATGCTAC 300  
DB 481 CCCCTGACACACACCCCTGAGACATCCGTTAGAGATGATGAGCTGCTGATGCTAC 540  
QY 301 ACAGATGTGAACGTCGCTGGAATGACAGAAATTTTCATGGGTGACTTCAATGCT 360  
DB 541 ACAGATGTGAACGTCGCTGGAAGGCGGAGAAATTTTCATGGGTGACTTCAATGCT 600  
QY 361 GGCTGAGCTACCTCCCAAGAGGCTGGAAGGACATCCGCTGAGAGGAGGACCCCAAG 420  
DB 601 GGCTGAGCTACCTCCCAAGAGGCTGGAAGGACATCCGCTGAGAGGACCCCAAG 660  
QY 421 TTGCTTTGGCTGATCGGGGACCAAGAGGACACCGCTCAAGAAAGACAAATCGCC 480  
DB 661 TTGCTTTGGCTGATCGGGGACCAAGAGGACACCGCTCAAGAAAGACAAATCGGCA 720  
QY 481 TATGACAGATCGTGTCTTACAGAGCAAAATATTTGTCATCTGTGTGCTCTCAATCAAC 540  
DB 721 TATGACAGATCGTGTCTTACAGAGCAAAATATTTGTCATCTGTGTGCTCTCAATCAAC 780  
QY 541 CTGCTCTTTGATTTCCAGAAAGCTTACAGGTTGCTGTAATCGAAGGCCCTGG 592  
DB 781 AGTGTGTTGACTTCCAGAAAGCTTACAGGTTGCTGTAATCGAAGGCCCTGG 832

RESULT 4  
LOCUS AX410746 1023 bp DNA linear PAT 14-JUN-2002  
DEFINITION Sequence 3393 from Patent WO0229103.  
ACCESSION AX410746  
VERSION AX410746.1 GI:21443451

SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1  
AUTHORS Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.  
TITLE Gene expression profiles in liver cancer  
JOURNAL Patent: WO 0229103-A 3393 11-APR-2002;  
GENE LOGIC INC (US)

FEATURES  
source 1..1023  
Location/Qualifiers  
1..1023  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/note="EMBL/GenBank Accession No. U56814"

BASE COUNT 312 a 244 c 240 g 227 t  
ORIGIN

Query Match 80.5%; Score 476.8; DB 6; Length 1023;  
Best Local Similarity 87.8%; Pred. No. 1.7e-119;  
Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 1 GAAACCTAAACGGAATTCAGAAAGGCAATACATACATATGATTTAGCTCCG 60  
DB 241 GAAACCTAAACGGAATTCAGAAAGGCAATACATACATATGATTTAGCTCCG 300  
QY 61 CTGGAAGAAACATATATAAGAAACATATGCTTTCTCTATAAGAAAGCTAGTGTCT 120  
DB 301 CTGGAAGAAACATATATAAGAAACATATGCTTTCTCTATAAGAAAGCTAGTGTCT 360  
QY 121 GTAAACAAAGCTACTCTACACAGACTATCAGGCTGAGAGCCAGATGTTTTCCAGG 180  
DB 361 GTGAAGAGGATTAATCACTACATGACTATCAGGATGAGACCCAGATGTTTTCCAGG 420  
QY 181 GAACCTTTGTGTCTGTTCCAGTACACCTACACCGCTGCAAGGCTCGTGAATGTC 240  
DB 421 GAACCTTTGTGTCTGTTCCAGTACACCTACACCTGCTGTCAAGACTTCTGATTTATC 480  
QY 241 CCCCTGACACACACCCCTGAGACATCCGTTAGAGATGATGAGCTGCTGATGCTAC 300  
DB 481 CCCCTGACACACACCCCTGAGACATCCGTTAGAGATGATGAGCTGCTGATGCTAC 540  
QY 301 ACAGATGTGAACGTCGCTGGAATGACAGAAATTTTCATGGGTGACTTCAATGCT 360  
DB 541 ACAGATGTGAACGTCGCTGGAAGGCGGAGAAATTTTCATGGGTGACTTCAATGCT 600  
QY 361 GGCTGAGCTACCTCCCAAGAGGCTGGAAGGACATCCGCTGAGAGGAGGACCCCAAG 420  
DB 601 GGCTGAGCTACCTCCCAAGAGGCTGGAAGGACATCCGCTGAGAGGAGGACCCCAAG 660  
QY 421 TTGCTTTGGCTGATCGGGGACCAAGAGGACACCGCTCAAGAAAGACAAATCGCC 480  
DB 661 TTGCTTTGGCTGATCGGGGACCAAGAGGACACCGCTCAAGAAAGACAAATCGTGA 720  
QY 481 TATGACAGATCGTGTCTTACAGAGCAAAATATTTGTCATCTGTGTGCTCTCAATCAAC 540  
DB 721 TATGACAGATCGTGTCTTACAGAGCAAAATATTTGTCATCTGTGTGCTCTCAATCAAC 780  
QY 541 CTGCTCTTTGATTTCCAGAAAGCTTACAGGTTGCTGTAATCGAAGGCCCTGG 592  
DB 781 AGTGTGTTGACTTCCAGAAAGCTTACAGGTTGCTGTAATCGAAGGCCCTGG 832

RESULT 5  
LOCUS HSU56814 1023 bp mRNA linear PRI 24-JUL-1997  
DEFINITION Human DNase1-like III protein (DNAS1L3) mRNA, complete cds.  
ACCESSION U56814  
VERSION U56814.1 GI:1399718

SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1023)  
AUTHORS Rodriguez, A.M., Rodin, D., Nomura, H., Morton, C.C., Weremowicz, S. and  
Schneider, M.C.  
TITLE Identification, localization, and expression of two novel human  
genes similar to deoxyribonuclease I  
JOURNAL Genomics 42 (3), 507-513 (1997)

REFERENCE 2 (bases 1 to 1023)  
AUTHORS Schneider, M.C. and Rodriguez, A.  
TITLE Direct Submission  
JOURNAL Submitted (25-APR-1996) M.C. Schneider, Renal Division, Brigham and  
Women's Hospital, 75 Francis Street, Boston, MA 02115, USA

FEATURES  
source 1..1023  
Location/Qualifiers  
1..1023  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

/clone\_lib="Stratagene liver (#937224)"  
/notes="corresponds to EST clones 82269 (GenBank Accession Number T6985, T69063), 82738 (GenBank Accession Number T73558, T73653), 78422 (GenBank Accession Number T61400, T61368)"  
1. .1023  
/gene="DNAS1L3"  
25. .942  
/gene="DNAS1L3"  
/codon\_start=1  
/product="DNase1-Like III protein"  
/protein\_id="AAB63967.1"  
/db\_xref="GI:1399713"  
/translation="MSRELAPLILLLLSHLSALAMRCSFNVSFGESKQEDKNAMDY IVKVKRCDIILVMEIKDSNNRRCIPILMEKLNRRRGITYNYVSSRLGRNTYKEQY AFLYKELVSKVSYHYDQGDADVFSREPVVWFQSPHTAVKDFVLIPLHTTPT SVKEIDELVEVYTDVKHRWKAENFIMGDFNAGCSYVKKAKNIRLRTDPRFVWLIG DQEDTTVKKSTNCAYDRIVLRGQEIYSSVVPKNSVDFQKAYKLTTEBALDVSDFHP VEFKQSSRAFTNSKSVTLRKTKSKRS"  
BASE COUNT 312 a 244 c 240 g 227 t  
ORIGIN

Query Match 80.5%; Score 476.8; DB 9; Length 1023;  
Best Local Similarity 87.8%; Pred. No. 1.7e-119;  
Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;  
Qy 1 GAGAGCTAAACGGAATTCAGAAAGGCATACATACAACTATGTAGTCTTCGC 60  
Db 241 GAGAGCTGAACAGAAATTCAGAGAGGCATACGTCACAACTATGTAGTCTCGG 300  
Qy 61 CTTGGAAGAACACATATAAGAACAGTATGCTTCTCTTAAGAAAGCTAGTCTCT 120  
Db 301 CTTGGAAGAACACATATAAGAACAAATATGCTTCTCTCAAGGAAAGCTGGTCT 360  
Qy 121 GTAAACAAAGCTACCTACACAGCTATCAGGCTGAGAGCAGAGTGTGTTTCCAGG 180  
Db 361 GTGAGAGGAGTTATCATCATGACTATCAGATGAGAGCAGAGTGTGTTTCCAGG 420  
Qy 181 GAACCTTTGTGTTGTGTTTCAGTCACTCCCTACACCGTGTCAAGGACTTCGTGATGTC 240  
Db 421 GAGCCCTTTGTGTTGTGTTTCCAATCTCCACACACTGCTGTCAAGAGCTTCGTGATATC 480  
Qy 241 CCCGTGACACACCCCTGAGACATCCGTTAGAGAGATTTGATGCTGGCTGATGTTCTAC 300  
Db 481 CCCGTGACACACCCCTGAGACATCCGTTAGAGATCGATGATGTTGTTGAGTCTAC 540  
Qy 301 ACAGATGGAACCGTCGCTGAATGCAGAGAAATTTCAATTTTCATGGTGTGACTTCAATGCT 360  
Db 541 ACGACGTGAACACCGCTGAGAGCGGAGAAATTTCAATTTTCATGGTGTGACTTCAATGCC 600  
Qy 361 GGCTGACGTACGTCCTCCCAAGAGCCCTGGAAGGACATCCGCTGAGAGCGGACCCCAAG 420  
Db 601 GGCTGACGTACGTCCTCCCAAGAGCCCTGGAAGAACATCCGCTTGAAGACTGACCCCAAG 660  
Qy 421 TTGCTTTGGCTGATCGGGGACCAAGAGACACACCGCTCAAGAGAGACACAACTGGCC 480  
Db 661 TTGCTTTGGCTGATCGGGGACCAAGAGACACACCGCTCAAGAGAGACACAACTGGCA 720  
Qy 481 TATGACAGATCGCTTTAGAGGACAAATATTTGCAACTCTGTTGTTGCTTCAATCAAAAC 540  
Db 721 TATGACAGATGTTGCTTTAGAGGACAAATGCTGACTTCTGTTTCCCAAGTCAAC 780  
Qy 541 CTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTCG 592  
Db 781 AGTGTGTTTGTGCTTCCAGAAAGCTTACAGCTGACTGAAGAGGAGGCCCTCG 832

RESULT 6  
BC015831  
LOCUS  
DEFINITION Homo sapiens, deoxyribonuclease I-like 3, clone MGC:27146 IMAGE:4723725, mRNA, complete cds.  
ACCESSION BC015831

BC015831.1 GI:16198370  
MGC.  
Homo sapiens.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1067)  
Strausberg,R.  
Direct Submission  
Submitted (15-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: http://mgc.nci.nih.gov  
Contact: MGC help desk  
Email: cgabbs@mail.nih.gov  
Tissue Procurement: CLONTECH  
cDNA Library Preparation: CLONTECH Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Institute for Systems Biology  
http://www.systemsbio.org  
Contact: amadan@systemsbiology.org  
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia Greene, Mark Kettman and Anuradha Madan  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAL.Plate: 38 Row: p Column: 7  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4826697.  
Location/Qualifiers  
1. .1067  
/organism="Homo sapiens"  
/db\_xref="LocusID:1776"  
/db\_xref="taxon:9606"  
/clone="MGC:27146 IMAGE:4723725"  
/tissue\_type="Liver"  
/clone\_lib="NIH\_MGC\_76"  
/lab\_host="DH10B"  
/note="Vector: pDNR-LIB"  
68. .985  
/codon\_start=1  
/product="deoxyribonuclease I-like 3"  
/protein\_id="AAH15831.1"  
/db\_xref="GI:16198371"  
/translation="MSRELAPLILLLLSHLSALAMRCSFNVSFGESKQEDKNAMDY IVKVKRCDIILVMEIKDSNNRRCIPILMEKLNRRRGITYNYVSSRLGRNTYKEQY AFLYKELVSKVSYHYDQGDADVFSREPVVWFQSPHTAVKDFVLIPLHTTPT SVKEIDELVEVYTDVKHRWKAENFIMGDFNAGCSYVKKAKNIRLRTDPRFVWLIG DQEDTTVKKSTNCAYDRIVLRGQEIYSSVVPKNSVDFQKAYKLTTEBALDVSDFHP VEFKQSSRAFTNSKSVTLRKTKSKRS"  
BASE COUNT 327 a 256 c 249 g 235 t  
ORIGIN

Query Match 80.5%; Score 476.8; DB 9; Length 1067;  
Best Local Similarity 87.8%; Pred. No. 1.7e-119;  
Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;  
Qy 1 GAGAGCTAAACGGAATTCAGAAAGGCATACATACAACTATGTAGTCTTCGC 60  
Db 284 GAGAGCTGAACAGAAATTCAGAGAGGCATACGTCACAACTATGTAGTCTCGG 343  
Qy 61 CTTGGAAGAACACATATAAGAACAGTATGCTTCTCTTAAGAAAGCTAGTCTCT 120  
Db 344 CTTGGAAGAACACATATAAGAACAAATATGCTTCTCTCAAGGAAAGCTGGTCT 403  
Qy 121 GTAAACAAAGCTACCTTACACGACTATCAGCTGGAGACGAGATGTTTCCAGG 180  
Db 404 GTGAGAGGAGTATATCACTACCATGACTATCAGATGAGAGCAGCATGTGTTCCAGG 463  
Qy 181 GAACCTTTGTGTTGTGTTTCCAGTCAACCTACACCGCTGTCAGGACTTCGTGATGTC 240  
Db 464 GAGCCCTTTGTGTTGTGTTTCCAATCTCCCAACACTCTCCCAAGACTTCGTGATATC 523



```
QY 241 CCCCTGACACACCCCTGAGACATCCCTTAGAGATGTAGTGTGCTGATGTCTAC 300
DB 524 CCCCTGACACACCCCTGAGACATCCCTTAGAGATGTAGTGTGCTGATGTCTAC 583
QY 301 ACAGATGTGAACCTGCTGATGAGACAGATTTTCAATTTTCATGGGTGACTTCAATGCT 360
DB 584 ACCGACGTGAACACCCCTGAGAGCGAGAAATTTTCATTTTCATGGGTGACTTCAATGCC 643
QY 361 GGCTGACGCTACGCTCCCAAGAAAGCCTGGAAGGACATCCGCTGAGAGAGAGACCCCAAG 420
DB 644 GGCTGACGCTACGCTCCCAAGAAAGCCTGGAAGGACATCCGCTGAGAGACTGACCCCAAG 703
QY 421 TTGCTTTGGTGTATCGGGGACCAAGAGACACACGCTCAAGAAAGACAACTGCGCC 480
DB 704 TTGCTTTGGTGTATCGGGGACCAAGAGACACACGCTCAAGAAAGACAACTGCGCA 763
QY 481 TATGACAGGATCGTCTTAAAGCAAAATTTGTCAATCTGTGTCTCTCAATCAAC 540
DB 764 TATGACAGGATCGTCTTAAAGCAAAATCGTCAATCTGTGTCTCTCAATCAAC 823
QY 541 CTCGCTTTGATTTCCAGAAAGCTTACAGGTTCTGAATCGAAGGCCCTGG 592
DB 824 AGCTTTTGAATTTCCAGAAAGCTTACAGGTTCTGAATCGAAGGCCCTGG 875

RESULT 7
AF047354 1079 bp mRNA linear PRI 16-SEP-1998
LOCUS Homo sapiens liver and spleen DNase precursor (LSD) mRNA, complete
DEFINITION cds.
ACCESSION AF047354
VERSION AF047354.1 GI:2905785
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1079)
Baker, W.F., Pan, C.Q., Spencer, S.A., Ryan, A.M., Lazarus, R.A. and
Baker, K.P.
Cloning and characterization of an actin-resistant DNase I-like
endonuclease secreted by macrophages
Gene 215 (2), 291-301 (1998)
9714828
2 (bases 1 to 1079)
Baron, W.F., Pan, C.Q., Spencer, S.A., Ryan, A.M., Lazarus, R.A. and
Baker, K.P.
Direct Submission
Submitted (09-FEB-1998) Molecular Biology, Genentech Inc., 1 DNA
Way, South San Francisco, CA 94080, USA
FEATURES
source
1..1079
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..1079
/gene="LSD"
71..988
/gene="LSD"
/notes="actin resistant activity; macrophage specific DNase
I; LS-DNase"
/codon_start=1
/product="liver and spleen DNase precursor"
/protein_id="AAC35752.1"
/db_xref="GI:2905786"
/translacion="MSRELAPLLLLLSHSLAMRIGSFNVSFGESKQEDKAMDV
IKVIRKCDIILWMEIKDSNNRILPLMKLRNRRGRTTYVVISRLGRYKXEOY
ALYKELVSVKSYHYHDYODGADADVEREPFVWFSQSPHTAVKDFVILPHTTRET
SVKEIDELVEVYTVDKHRKAKENFIEMGFPNAGCSYPRKAKNIRLRTPDFWILG
DOEDTTVKSTNCATDRILVRCOEIVSSVPSNSVDFQAKYKLTLEBALDVSHPF
VEFKQSSRAFTNSKSVTLRKTKRSKRS"
71..130
sig_peptide
```

```
/gene="LSD"
/evidence=experimental
mat_peptide 131..985
/gene="LSD"
/product="liver and spleen DNase"
BASE COUNT 330 a 260 c 252 g 237 t
ORIGIN
Query Match 80.5%; Score 476.8; DB 9; Length 1079;
Best Local Similarity 87.8%; Pred. No. 1,7e-119;
Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
QY 1 GAGAACTAAACGAAATTCAGAAAGGCAATACATCAATGATGATTACTCTGCG 60
DB 287 GAGAGCTGAACAGAAATTCAGAAAGGCAATACATCAATGATGATTACTCTGCG 346
QY 61 CTTGGAAGAAACATATTAAGAAACGATATGCTTTCTCTATTAAGAAAGCTAGTCT 120
DB 347 CTTGGAAGAAACATATTAAGAAACATATGCTTTCTCTATTAAGAAAGCTAGTCT 406
QY 121 GTAAACAAAGCTACCTTACACGACTATCAGCTGGAGAGCAGATGTTTCCAGG 180
DB 407 GTAAAGAGATTAATCACTACATGACTATCAGATGAGAGACGATGTTTCCAGG 466
QY 181 GAACCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
DB 467 GAGCCCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 526
QY 241 CCCCTGACACACCCCTGAGACATCCCTTAGAGATGTAGTGTGCTGATGTCTAC 300
DB 527 CCCCTGACACACCCCTGAGACATCCCTTAGAGATGTAGTGTGCTGATGTCTAC 586
QY 301 ACAGATGTGAACCTGCTGAGATGAGAAATTTTATTTTATGCTGACTTCAATGCT 360
DB 587 ACAGATGTGAACCTGCTGAGATGAGAAATTTTATTTTATGCTGACTTCAATGCT 646
QY 361 GGCTGACGCTACGCTCCCAAGAAAGCCTGGAAGGACATCCGCTGAGAGAGAGAGAG 420
DB 647 GGCTGACGCTACGCTCCCAAGAAAGCCTGGAAGGACATCCGCTGAGAGAGAGAGAG 706
QY 421 TTGCTTTGGTGTATCGGGGACCAAGAGACACACGCTCAAGAAAGCAAACTGCGCC 480
DB 707 TTGCTTTGGTGTATCGGGGACCAAGAGACACACGCTCAAGAAAGCAAACTGCGCA 766
QY 481 TATGACAGGATCGTCTTAAAGCAAAATTTGTCAATCTGTGTCTCTCAATCAAC 540
DB 767 TATGACAGGATCGTCTTAAAGCAAAATCGTCAATCTGTGTCTCTCAATCAAC 826
QY 541 CTCGCTTTGATTTCCAGAAAGCTTACAGGTTCTGAATCGAAGGCCCTGG 592
DB 827 AGCTTTTGAATTTCCAGAAAGCTTACAGGTTCTGAATCGAAGGCCCTGG 878

RESULT 8
HSU75744 1108 bp mRNA linear PRI 17-JUN-1998
LOCUS Homo sapiens DNase gamma mRNA, complete cds.
DEFINITION HSU75744
ACCESSION U75744
VERSION U75744.1 GI:3236319
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1108)
Shiohawa, D., Hiraizumi, M., and Tanuma, S.
CDNA cloning of human DNase gamma: chromosomal localization of its
gene and enzymatic properties of recombinant protein
Apoptosis 3 (2), 89-95 (1998)
2 (bases 1 to 1108)
Shiohawa, D., and Tanuma, S.
Direct Submission
Submitted (23-OCT-1996) Biochemistry, Science Univ. of Tokyo,
```

```

FEATURES             Shinjuku-ku Ichigaya Funagawaracho, Tokyo 162, Japan
source               Location/Qualifiers
1..1108
/organism="Homo sapiens"
/db_xref="taxon:9606"
93..1010
/functions="apoptosis"
/function="deoxyribonuclease"
/codon_start=1
/product="DNase gamma"
/protein_id="AAC323652.1"
/db_xref="GI:3236320"
/translators="MSRELAPL11LHLSALAMRCSFNVRSGESKODKNDV
IVKVKCDIILWEIKDNNRCPILMEKLNRSRGITYNVIVSRGRKTYKEQY
AFLYKELSVKSYHYHDYQGDADVFSPEFVVFQSPHTAVKDFVILPLHTPET
SVKIDELVEYTDVRRKAEFIEMGDFNAGSVYPKAKWNIKRLTDPFRFVLIG
DOEDTVKKSNCAYDRIVLRQGEIVSVVYKNSVDFQKAYKLTETEEALDVSDHP
VEPKLQSSRAFTNSKSVTLRKKTKSKRS"
BASE COUNT        329 a 261 c 271 g 247 t
ORIGIN
Query Match      80.3%; Score 475.2; DB 9; Length 1108;
Best Local Similarity 87.7%; Pred. No. 4.7e-119;
Matches 519; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
Qy 1 GAGAGCTAAACGGAAATCAAGAAAGGCATACATACATATGATAGCTCTCGC 60
Db 309 GAGAGCTGAACAGAAATCAAGGAGAGGCATACATATGATAGCTCTCGG 368
Qy 61 CTGGAAGAAACACATATAAGAACAGTATGCTTTCTCTATAAGAAAGCTAGTGCT 120
Db 369 CTGGAAGAAACACATATAAGAACATATGCTTTCTCTACAGGAAGAACTGCTGCT 428
Qy 121 GTAAACAAAGCTACCTTACCAGACTATCAGGCTGGAGCGCAGATGTTTTCAGG 180
Db 429 GTGAAGAGGAGTTATCACTACCATGACTATCAGGATGAGAGCGCAGATGTTTTCAGG 488
Qy 181 GAACCTTTGTGCTGCTTCCAGTCACTACCCGTACACCGCTGTCAAGGACTTCTGATGTC 240
Db 489 GAGCCCTTTGTGCTGCTTCCAAATCTCCCACTGCTGTCAAGAGCTTCTGATATC 548
Qy 241 CCCCTGCACACCCCTGAGACATCCGTTAGAGAGATGATGCTGGCTGATGCTTAC 300
Db 549 CCCCTGCACACCCCGAGACATCCGTTAAGAGATCGATGATGTTGGTTGAGTCTAC 608
Qy 301 ACAGATGGAACCGTCTGGAATGACAGAAATTTCAATTTTCATGGTGACTTCAATGCT 360
Db 609 ACGGACGTGAACACCGCTGGAAGCGGAGAAATTTCAATTTTCATGGTGACTTCAATGCC 668
Qy 361 GGCTGCAGCTAGTCCCAAGAGCGCTGGAAGAGCATCCGCTGAGGACGAGCCCAAG 420
Db 669 GGCTGCAGCTAGTCCCAAGAGCGCTGGAAGAACATCCGCTTGAAGACTGAGCCCAAG 728
Qy 421 TTGCTTTGCTGATCGGGGACCAAGAGACACCGCTCAAGAGAGACACAACTGGGCC 480
Db 729 TTTGTTGGCTGATCGGGGACCAAGAGACACCGCTGGAAGAGAGACCAACTGTGCA 788
Qy 481 TATGACAGGATCGCTTTAGAGGACAAATATTTGCAACTCTGTTGGTGCTTCAATCAAAAC 540
Db 789 TATGACAGGATGCTTTAGAGGACAAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAAAC 848
Qy 541 CTCGCTTTGATTTCCAGAAAGCTTACAGGTTGCTGAAATCGAAGGCGCTCG 592
Db 849 AGTGTGTTTGACTTCCAGAAAGCTTACAGCTGACTGAAGAGGAGGCGCTCG 900
RESULT 9
AF039852 1331 bp mRNA linear ROD 07-AUG-1998
LOCUS      Rattus norvegicus DNaseI mRNA, complete cds.
DEFINITION AF039852
ACCESSION  AF039852
VERSION    AF039852.1 GI:3395771
KEYWORDS

SOURCE
ORGANISM
Rattus norvegicus.
Rattus norvegicus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 1331)
AUTHORS
Liu,Q.Y., Pandey,S., Singh,R.K., Lin,W., Ribecco,M.,
Borowy-Borowski,H., Smith,B., LeBlanc,J., Walker,P.R. and
Sikorska,M.
TITLE
DNaseI: a rat DNaseI-like gene coding for a constitutively
expressed chromatin-bound endonuclease
JOURNAL
Biochemistry 37 (28), 10134-10143 (1998)
MEDLINE
9832537
PUBMED
9665719
2 (bases 1 to 1331)
AUTHORS
Liu,Q.Y., Singh,R.K., Lin,W. and Sikorska,M.
TITLE
Direct Submission
JOURNAL
Submitted (24-DEC-1997) Institute for Biological Sciences, National
Research Council, 1200 Montreal Road, Bldg. M-54, Box 4, Ottawa, ON
K1A 0R6, Canada
FEATURES
Location/Qualifiers
1..1331
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/tissue_type="liver"
153..1085
/function="apoptotic endonuclease"
/note="similar to rat DNase I; chromatin-bound"
/codon_start=1
/product="DNaseI"
/protein_id="AAC28937.1"
/db_xref="GI:3395772"
/translators="MSLYPASPYLASLLFLLALHGALSRLCSFNVRSGESKKNH
NAMDIIKIKRCDLILMEIKDNNRCPILMEKLNRSRGITYNVIVSRGRNT
YEQVAFLYKELSVKAYLVHDYQGDVDFSPFVVFQAPPTAAKDFVILPLH
TTPESVKEIDELADVYTDVRRKAEFIEMGDFNAGSVYPKAKWNIKRLTDPNF
VMLQGEDTVKKSNCAYDRIVLRQGEIVSVVYKNSVDFQKAYKLTETEEALDV
SDHPVFEPKLSRAFTNSKSVTLRKKTKSKRS"
BASE COUNT        351 a 333 c 320 g 327 t
ORIGIN
Query Match      77.6%; Score 459.2; DB 10; Length 1331;
Best Local Similarity 86.0%; Pred. No. 1.1e-114;
Matches 509; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
Qy 1 GAGAGCTAAACGGAAATTCAGAAAGGCATACATACATATGATAGCTCTCGC 60
Db 384 GAGAGCTGAATGGAATCTCAGAGAGAGCAGACATACATACATGATAGCTCTCGG 443
Qy 61 CTTGGAAGAAACACATATAAGAACAGTATGCTTTCTTATAAGAAAGCTAGTGCT 120
Db 444 CTTGGAAGAAACACATATAAGAACAGTATGCTTCTCTACAAGGAGAGCTGCTCT 503
Qy 121 GTAAACAAAGCTACCTTACCAGACTATCAGGCTGGAGCGCAGATGTTTTCAGG 180
Db 504 GTGAAGGCAAAATACCTCTACCATGACTATCAGGATGGAGACACACGCTGTTTCAGG 563
Qy 181 GAACCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 564 GAGCCCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 623
Qy 241 CCCCTGCACACCCCTGAGACATCGGTTAGAGAGATGATGAGCTGGCTGATGCTAC 300
Db 624 CCCCTGCACAACTCCTGAAACCTCGTTTAAAGAGATAGATGAGCTGGCTGCTAC 683
Qy 301 ACAGATGTAAGCTGCTGTAATGCAGAGAAATTCATTTTCATGGTGACTTCAATGCT 360
Db 684 ACGATGTGAGAGACCATGTAAGAGGAGAGAAATTCATTTTCATGGTGATTTCAATGCT 743
Qy 361 GGCTGCAGCTACGTCCTCCCAAGAGGCGCTGGAAGGACATCCGCTGAGGACGAGCCCAAG 420
Db 744 GGCTGCAGCTACGTCCTCCCAAGAGGCGCTGGAAGAACATCCGTTTGAAGACAGACCCCAAC 803
```

QY	421	TTTCGTTGGCTATCGGGACCAAGAGACACACCGTCAAGAAAGACCAAACTGCGCC	480
DB	804	TTTGTGTGGCTATTTGGGACCAAGAGACACACCGTCAAGAAAGACCAAGCTGTGCC	863
QY	481	TATGACAGATCGTGCTTAGAGACAAAATATGTCAACTGTGGTGTCTCAATCAAC	540
DB	864	TATGACAGATGTGCTTGGCGGACAGAGATAGTCAACTGTGTGTTCCCGCTCAGT	923
QY	541	CTCGCTTTGATTTCCAGAAACTTACAGGTGTCTGAATCGAAGGCGCTTG	592
DB	924	GCGCTTTGACTTTCAGAAACTTATGAGTGTCTGAAAGAGAGCGCTTG	975
RESULT 10			
LOCUS	RNU75689	1417 bp	mRNA linear ROD 15-JUN-1998
DEFINITION	Rattus norvegicus DNaase gamma mRNA, complete cds.		
ACCESSION	U75689		
VERSION	U75689.1	GI:3220024	
KEYWORDS			
SOURCE	Rattus norvegicus.		
ORGANISM	Rattus norvegicus.		
REFERENCE			
AUTHORS	Shioikawa,D., Ohyama,H., Yamada,T., Takahashi,K. and Tanuma,S.		
TITLE	Identification of an endonuclease responsible for apoptosis in rat thymocytes		
JOURNAL	Eur. J. Biochem. 226 (1), 23-30 (1994)		
MEDLINE	95045594		
PUBMED	7957253		
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			

Db	434	GAGAAAGCTGAATGGAATCTCAGAAAGACAGCATCAACTAGCTGATTAGCTTCGG	493
Qy	61	CTTGGAAAGAAACACATATTAAGAACAGATAGCTTCTCTATAAAGAAAGCTAGTGTCT	120
Db	494	CTTGGAAAGAAACACATATTAAGAAAGATAGCTTCTCTATAAAGAAAGCTAGTGTCT	553
Qy	121	GTAAGAACAAAGCTACTCTTACACAGCATATCAGGCTGGAGACGAGATGTGTTTCAGG	180
Db	554	GTTGAAGGCAAAATACCTTACCATATGACTATACAGATGGAGACACAGCTGTTCAGG	613
Qy	181	GAACCTTTGT	240
Db	614	GAGCCCTTTGT	673
Qy	241	CCCTGTCACACACCCCTGAGACATCCGTTAGAGATTTAGTACCTGCTGATCTAC	300
Db	674	CCCTTGCACACACCTCTGAAACCTCCCTTTAAGAGATAGATGAGCTGCTGACCTTAC	733
Qy	301	ACAGATGGAACGCGGTGGAAATGACAGAAATTTCAATTTTCAATGGTGAATTTCAATGCT	360
Db	734	ACGGATGTGAAGAACGATGGAAGGACAGAAATTTCAATTTTCAATGGTGAATTTCAATGCT	793
Qy	361	GGCTGACAGCTACGTCCTCCCAAGAGGCTGGAAGACATCCGCTGAGACGGAACCCCAAG	420
Db	794	GGCTGACAGCTACGTCCTCCCAAGAGGCTGGAAGACATCCGTTGAGACAGACCCCAAC	853
Qy	421	TTTCGTTGGCTGATTCGGGACCAAGAGACACACGCTCAAGAAAGACAACTGCGCC	480
Db	854	TTTCGTTGGCTGATTCGGGACCAAGAGACACACGCTCAAGAAAGACCACTGCTGCC	913
Qy	481	TATGACAGGATCGTGTCTTAGAGACAAATATTTGTCATCTGTGTGTCTTCAATCAAC	540
Db	914	TATGACAGGATCGTGTCTTAGAGACAAATATTTGTCATCTGTGTGTCTTCAATCAAC	973
Qy	541	CTCGCTTGTGATTTCCAGAAAGCTTACAGGTGTGTGTGATGGAAGGCTCGG	592
Db	974	GCGCTTTGTGATTTCCAGAAAGCTTATGAGATGTGTGTGGAAGGAGGCTCGG	1025
RESULT 11			
AR047845			
LOCUS	AR047845	1208 bp	DNA
DEFINITION	Sequence 2 from patent US 5821103.	linear	PAT 29-SEP-1999
ACCESSION	AR047845		
VERSION	AR047845.1	GI:5970188	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 1208)		
TITLE	Tanuma, S.-i.		
JOURNAL	Deoxyribonuclease		
FEATURES	Patent: US 5821103-A 2 13-OCT-1998;		
	Location/Qualifiers		
	1..1208		
	/organism="unknown"		
BASE COUNT	319 a	304 c	283 g
ORIGIN		302 t	
Query Match			
Best Local Similarity 76.5%; Score 452.8; DB 6; Length 1208;			
Matches 505; Conservativity 85.3%; Pred. No. 6,4e-113;			
Matches 505; Conservativity 0; Mismatches 87; Indels 0; Gaps 0;			
Qy	1	GAGAGCTTAACGGAATTTCAAGAAAGCATTAACATACATATGTGATTAGCTTTCGC	60
Db	243	GAGAGCTGAATGGAATCTACGAAAGACGACGATTAACATACGATTAAGCTTCGCG	302
Qy	61	CTTGAAGAAACATATTAAGAAACAGTATGCTTCTCTATAAAGAAAGCTAGTGTCT	120
Db	303	CTTGAAGAAACATATTAAGAAACAGTATGCTTCTCTATAAAGAAAGCTAGTGTCT	362
Qy	121	GTAAGAACAAAGCTACTCTTACACAGCATATCAGGCTGGAGACGAGATGTGTTTCAGG	180

```
Db 363 GTGAGGCAAAATACCTCTACCATGACTATCAGGATGGAGACACAGAGCTGTTTTCAGG 422
Qy 181 GAACCCCTTTGGTGTGGTTCAGGTCACCCCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
Db 423 GAGCCCTTTGGTGTGGTTCAGGCGCCCTTCACTGCTGCGCAAGGACTTCGTGATTGTC 482
Qy 241 CCCCTGCACACCCCTGAGACATCCGTTAGAGATTCGATGAGCTGGCTGATGTTCTAC 300
Db 483 CCCCTTGACACAACTCTCTGAAACCTCCGTTAAAGAGATAGATGAGCTGGCTGATGTTCTAC 542
Qy 301 ACAGATGGAACCTGCTGGAATGCAGAGAAATTCATTTTCATGGGTGACTTCAATGCT 360
Db 543 ACGGATGTTAGAGACATGGAAGGAGAGATTTTCATCTTCATGGGTGATTTCAATGCT 602
Qy 361 GGCTGCAGCTACGTCCTCCCAAGAGCCCTGGAAGGACATCCGCTGAGACCGACCCCAAG 420
Db 603 GGCTGCAGCTACGTCCTCCCAAGAGCCCTGGAAGAACATCCGTTTGGAGACAGACCCCAAC 662
Qy 421 TTCGTTGGCTGATCGGGGACCAAGAGACACACCGCTCAAGAGAGACACAACTGCGCC 480
Db 663 TTTGTTGGCTGATTTGGGGACCAAGAGACACACCGCTCAAGAGAGACACAGCTGTGCC 722
Qy 481 TATGACAGATCGCTGTTAGAGGACAAATATTTGTCACCTCTGTTGGTCCCTCAATCAAAC 540
Db 723 TATGACAGATGTTGCTTCGGGACCAAGAGACACACCGCTCAAGAGAGACACAGCTGTGCC 782
Qy 541 CTCGCTTTGATTTTCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
Db 783 GGCGCTTTGACTTTCAGAAAGCTTATGAGTTGTCTGAAGAGGAGGCCCTGG 834

RESULT 12
E11687 1208 bp RNA linear PAT 29-SEP-1997
LOCUS cDNA encoding novel deoxyribonuclease (DNase) gamma which cut
DEFINITION specifically linker site of chromatin DNA.
ACCESSION E11687
VERSION E11687.1 GI:22025323
KEYWORDS JP 1996187079-A/1.
SOURCE Rattus rattus.
ORGANISM Rattus rattus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1208)
AUTHORS Tanuma, Y.
TITLE NEW DEOXYRIBONUCLEASE
JOURNAL Patent: JP 1996187079-A 1 23-JUL-1996;
TANUMA YASUKAZU
COMMENT OS Rattus rattus (rat)
PN JP 1996187079-A/1
PD 23-JUL-1996
PF 06-SEP-1995 JP 1995255647
PR 06-SEP-1994 JP 94P 239518
PT TANUMA YASUKAZU
PC C12N9/16,C07K14/47,C07K16/40,C12N1/21,C12N15/09,(C12N1/21, PC
C12R1:19);
CC strandedness: Double;
CC topology: linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1. .1208
FT /organism='Rattus rattus'
FT /tissue type='thymus, Spleen' FT
FT /cell type='Cell nucleus'
FT 5'UTR 1. .11
FT CDS 12. .944
FT /product='Novel deoxyribonuclease (DNase) gamma
FT which cut
FT specifically linker site of chromatin DNA' FT
FT precursor_RNA 12. .86
```

```
FEATURES
source
Location/Qualifiers
1. .1208
/organism='Rattus rattus'
/db_xref='taxon:10117'
BASE COUNT 319 a 304 c 283 g 302 t
ORIGIN
Query Match 76.5%; Score 452.8; DB 6; Length 1208;
Best Local Similarity 85.3%; Pred. No. 6.4e-113;
Matches 505; Conservative 0; Mismatches 87; Indels 0; Gaps 0;
Qy 1 GAGAGCTAAACCGAATTCAGAAAGGCATACATACATATGTTAGTAGTCTCGC 60
Db 243 GAGAAGCTGAATGGAACCTCAGGAAGACGACATACATACATGTTAGTCTCGG 302
Qy 61 CTTGGAAGAAACACATATAAAGAACAGTATGCCCTTCTTATATAAAGAAAGTGTCT 120
Db 303 CTTGGAAGAAACACATATAAAGAACAGTATGCCCTTCTTACAAAGGAGAGTGTCT 362
Qy 121 GTAAACAAAGCTACCTCTACACGACTATCAGGCTGGAGACGAGATGTGTTCCAGG 180
Db 363 GTGAAGGCAAAATACCTCTACCATGACTATCAGGATGGAGACACAGACGTGTTTCAGG 422
Qy 181 GAACCCCTTTGGTGTGGTTCAGTCCAGTCCACCTACACCGCTGTCAAGGACTTCGTGATTGTC 240
Db 423 GAGCCCTTTGGTGTGGTTCAGGCGCCCTTCACTGCTGCGCAAGGACTTCGTGATTGTC 482
Qy 241 CCCCTGCACACCCCTGAGACATCCGTTAGAGAGATTCGATGAGCTGGCTGATGTTCTAC 300
Db 483 CCCTTGACACAACTCTCTGAAACCTCGTTAAAGAGATAGATGAGCTGGCTGATGTTCTAC 542
Qy 301 ACAGATGGAACCTGCTGGAATGCAGAGAAATTCATTTTCATGGGTGACTTCAATGCT 360
Db 543 ACGGATGTTAGAGACATGGAAGGAGAGATTTTCATCTTCATGGGTGATTTCAATGCT 602
Qy 361 GGCTGCAGCTACGTCCTCCCAAGAGCCCTGGAAGGACATCCGCTGAGACCGACCCCAAG 420
Db 603 GGCTGCAGCTACGTCCTCCCAAGAGCCCTGGAAGAACATCCGTTTGGAGACAGACCCCAAC 662
Qy 421 TTCGTTGGCTGATCGGGGACCAAGAGACACACCGCTCAAGAGAGACACAACTGCGCC 480
Db 663 TTTGTTGGCTGATTTGGGGACCAAGAGACACACCGCTCAAGAGAGACACAGCTGTGCC 722
Qy 481 TATGACAGATCGCTGTTAGAGGACAAATATTTGTCACCTCTGTTGGTCCCTCAATCAAAC 540
Db 723 TATGACAGATGTTGCTTCGGGACCAAGAGACACACCGCTCAAGAGAGACACAGCTGTGCC 782
Qy 541 CTCGCTTTGATTTTCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
Db 783 GGCGCTTTGACTTTCAGAAAGCTTATGAGTTGTCTGAAGAGGAGGCCCTGG 834

RESULT 13
MMU76110
LOCUS Mus musculus DNase gamma mRNA, complete cds.
DEFINITION
ACCESSION U76110
VERSION U76110.1 GI:4098207
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 941)
AUTHORS Shiohara, D., Hatanaka, T. and Tanuma, S.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-1996) Biochemistry, Science Univ. of Tokyo,
Shinjuku-ku ichigaya funagawaracho, Tokyo 162, Japan
FEATURES
source
Location/Qualifiers
1. .941
/organism='Mus musculus'
/strain='C57 black'
```

CDS

/db\_xref="taxon:10090"  
 7. .939  
 /function="apoptosis"  
 /note="deoxyribonuclease"  
 /codon\_start=1  
 /product="DNase gamma"  
 /protein\_id="A009222.1"  
 /db\_xref="GI:4098208"  
 /translation="MSLHPASPRSLALFILALHDTLALRLCSFNVRSFGASKKENH  
 EAMDIIVKIKRCDLILMEIKDSSNNICPMLEKLNRSRSTTYVVISRLGRNT  
 YKEOYAFVYKKEIKLVSKTKYHHDYODDPTVFSRPPVWFHSPTAYKDYIVLH  
 TTPERSVKEIDELVDYTVRSOMKTEFIIMGDFNAGCSYPPKKAMNIRLTDPKF  
 VWLIGDEDTVKSCTAYDRIVLGGEBIVNSVVRSSGVDFQKAYDLSEELADV  
 SHFPVEFLQSSRAFTNNRKSLSLKKRKKGNRS"

BASE COUNT 255 a 229 c 228 g 229 t

Query Match 73.5%; Score 435.2; DB 10; Length 941;  
 Best Local Similarity 83.4%; Pred. No. 4.1e-108;  
 Matches 494; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

1 GAGAGCTAAACGGAATTCAGAAAGGATACATACATGATTAGTCTCGC 60  
 238 GAGAGCTGAATGAAATTCAGAAAGGATACATACATGATTAGTCTCGA 297  
 61 CTTGGAAGAAACATATTAAGAAACAGTATGCTTCTCTATTAAGAAAGTATGCT 120  
 298 CTTGGAAGAAACGTAACAAGAGCATATGCTTCTCTACAAAGAGAGTGTGTCT 357  
 121 GTTAAACAAAGTACTCTACAGACTATCAGGCTGAGAGCGAGATGTGTTTCCAGG 180  
 358 GTGAAACAAATTAACATCACTGATCAGATGAGAGACAGACGCTGTTTCCAGG 417  
 181 GAACCTTTGTGCTGCTGCTCCAGTCACTACACCGCTTCAAGACTTCGCTGATGTC 240  
 418 GAACCTTTGTGCTGCTGCTCCAGTCACTACACCGCTTCAAGACTTCGCTGATGTC 477  
 241 CCCCTGACACCAACCCCTGAGATCCGTTAGAGAGATTGATGAGCTGATGCTTAC 300  
 478 CCTTGACACCAACCTCCGAGACCTCGTTAAAGATATGATGCTGCTGATGCTAC 537  
 301 ACAGATGTGAAGCTGCTGCTGAGATGAGAGATTTCTATGAGTGAATGCT 360  
 538 ACAGATGTGAAGCTGCTGCTGAGATGAGAGATTTCTATGAGTGAATGCT 597  
 361 GGCTGAGTACGCTCCCAAGAGGCTGGAAGACATCCGCTGAGAGAGGCCCAAG 420  
 598 GGCTGAGTACGCTCCCAAGAGGCTGGAAGACATCCGCTGAGAGAGGCCCAAG 657  
 421 TTGCTTGGCTGATCGGGGACCAAGAGACACCGGTCMAAGAGACCAACCTGCGC 480  
 658 TTGCTTGGCTGATCGGGGACCAAGAGACACCGGTCMAAGAGACCAACCTGCGC 717  
 481 TATGACAGATCGTCTTGAAGACAAATATTTGCACTCTGCTGCTCTCAATCAAC 540  
 718 TATGACAGATCGTCTTGAAGACAAATATTTGCACTCTGCTGCTCTCAATCAAC 777  
 541 CTCGCTTGTGATTTCCAGAAAGCTTACAGTTCTGGAATCGAAGGCCCTGG 592  
 778 GCGCTTGTGATTTCCAGAAAGCTTACAGTTCTGGAATCGAAGGCCCTGG 829

RESULT 14  
 AF047355 1124 bp mRNA linear ROD 16-SEP-1998  
 LOCUS AF047355 Mus musculus liver and spleen DNase precursor (LSD) mRNA, complete  
 DEFINITION cds.  
 ACCESSION AF047355  
 VERSION AF047355.1 GI:2905787  
 KEYWORDS  
 SOURCE Mus musculus.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
 1 (bases 1 to 1124)  
 AUTHORS Baron,W.F., Pan,C.Q., Spencer,S.A., Ryan,A.M., Lazarus,R.A. and Baker,K.P.  
 TITLE Cloning and characterization of an actin-resistant DNase I-like endonuclease secreted by macrophages  
 JOURNAL Gene 215 (2), 291-301 (1998)  
 MEDLINE 98382522  
 PUBMED 9714828  
 REFERENCE 2 (bases 1 to 1124)  
 AUTHORS Baron,W.F., Pan,C.Q., Spencer,S.A., Ryan,A.M., Lazarus,R.A. and Baker,K.P.  
 TITLE Direct Submission  
 JOURNAL Submitted (09-FEB-1998) Molecular Biology, Genentech Inc., 1 DNA Way, South San Francisco, CA 94080, USA

FEATURES  
 source  
 1. .1124  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 1. .1124  
 /gene="LSD"  
 173. .1105  
 /gene="LSD"  
 /note="actin resistant activity; macrophage specific DNase I; LS-DNase"  
 /product="liver and spleen DNase precursor"  
 /protein\_id="AAC35753.1"  
 /db\_xref="GI:2905787"  
 /translation="MSLHPASPRSLALFILALHDTLALRLCSFNVRSFGASKKENH  
 EAMDIIVKIKRCDLILMEIKDSSNNICPMLEKLNRSRSTTYVVISRLGRNT  
 YKEOYAFVYKKEIKLVSKTKYHHDYODDPTVFSRPPVWFHSPTAYKDYIVLH  
 TTPERSVKEIDELVDYTVRSOMKTEFIIMGDFNAGCSYPPKKAMNIRLTDPKF  
 VWLIGDEDTVKSCTAYDRIVLGGEBIVNSVVRSSGVDFQKAYDLSEELADV  
 SDHPVEFLQSSRAFTNNRKSLSLKKRKKGNRS"  
 173. .247  
 /gene="LSD"  
 /evidence="not\_experimental"  
 248. .1102  
 /gene="LSD"  
 /product="liver and spleen DNase"

mat\_peptide  
 BASE COUNT 299 a 273 c 273 g 279 t

Query Match 73.5%; Score 435.2; DB 10; Length 1124;  
 Best Local Similarity 83.4%; Pred. No. 4.1e-108;  
 Matches 494; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

1 GAGAGCTAAACGGAATTCAGAAAGGATACATACATGATTAGTCTCGC 60  
 404 GAGAGCTGAATGAAATTCAGAAAGGATACATACATGATTAGTCTCGA 463  
 61 CTTGGAAGAAACATATTAAGAAACAGTATGCTTCTCTATTAAGAAAGTATGCT 120  
 464 CTTGGAAGAAACAGTATTAAGAAAGCAGTATGCTTCTCTACAAAGAGTGTGTCT 523  
 121 GTTAAACAAAGTACTCTACAGACTATCAGGCTGAGAGCGAGATGTGTTTCCAGG 180  
 524 GTGAAACAAATTAACATCACTGATGAGATGAGAGACAGACGCTGTTTCCAGG 583  
 181 GAACCTTTGTGCTGCTGCTCCAGTCACTACACCGGTCMAAGAGACTTGTGATTGTC 240  
 584 GAACCTTTGTGCTGCTGCTCCAGTCACTACACCGGTCMAAGAGACTTGTGATTGTC 643  
 241 CCCCTGACACCAACCCCTGAGATCCGTTAGAGAGATTGATGAGCTGATGCTTAC 300  
 644 CCTTGACACCAACCTCCGAGACCTCCGTTAAAGATAGATGAGCTGCTGATGCTTAC 703  
 301 ACAGATGTGAAGCTGCTGCTGAGATGAGAGATTTCTATGAGTGAATGCT 360  
 704 ACAGATGTGAAGCTGCTGCTGAGATGAGAGATTTCTATGAGTGAATGCT 763  
 361 GGCTGAGTACGCTCCCAAGAGGCTGGAAGACATCCGCTGAGAGCGAACCCCAAG 420

```
||||| 764 GGCTGTAGCTATGTCCTCCCAAGAGCTGGCAGAACATTCGTTTGAGACGGACCCCAAG 823
||||| QY TTCTGTTGGCTGATCGGGGCAACAGAGACACACCGGTCAAGAGAGACACAACTGGGCC 480
||||| Db TTTGTTGGCTGATTTGGGACCAAGAGACACTACGGTCAAGAGAGATACAGCTGTGCC 883
||||| QY TATGACAGGATCGTCTTAGAGGACAAATATGTCAACTCTGTGTGGTCTCAATCAAAC 540
||||| Db TATGACAGGATTTGCTTTGTGGACAGAGATAGTCAACTCCGTGGTTCCTCCGTTCCAGT 943
||||| QY CTGCTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592
||||| Db GCGCTCTTTGATTTTCAGAAAGCTTATGACTTGTCTGAGGAGGAGGCCCTGG 995

RESULT 15
BC012671 2110 bp mRNA linear ROD 07-AUG-2002
LOCUS Mus musculus, Similar to deoxyribonuclease 1-like 3, clone
DEFINITION MGC:13854 IMAGE:4160709, mRNA, complete cds.
ACCESSION BC012671
KEYWORDS BC012671.1 GI:15215118
SOURCE MGC.
ORGANISM house mouse.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2110)
Strausberg, R.
Direct Submission
Submitted (15-AUG-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadanesystemsbio.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: INRAK Plate: 18 Row: 1 Column: 14
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6681204.
Location/Qualifiers
1..2110
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:13854 IMAGE:4160709"
/tissue_type="Liver, normal. 5 month old male mouse."
/clone_lib="NCI CGAP_L19"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
151..1083
/codon_start=1
/product="Similar to deoxyribonuclease 1-like 3"
/protein_id="AAH12671.1"
/db_xref="GI:15215118"
/translation="MSLHPASPRLASLLFLIALHDTLALRCSFNVRSPGASKENH
EAMDIIIVKIKRDLILLEIKDSNNICPMLMEKLNRSRTIINYIVISSRLGRNT
YKEQVAFVYKELSVKTKYHYDQDGTDFRPFVWFHSPFTAVKDFIVPLH
TPPETSVEIDELVDYTVDRSQWKTENFIEMGDFNAGCSYVPKKAWQINRLRDPKF
WLLIGDQEDTTVKKSTSCAYDRIVLCGOEIVNSVVPSSGVSDFQKAYLSEEDLDV
SDHPFVEFKLOSSRAFTNNRKSLSLKKKGNRS"
```

```
BASE COUNT 526 a 581 c 464 g 539 t
ORIGIN
Query Match 73.2%; Score 433.6; DB 10; Length 2110;
Best Local Similarity 83.3%; Pred. No. 1.2e-107;
Matches 493; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY 1 GAGAAGCTAAACGGAAATTCAGAAAGAGCATACACATACTATGTGATTAGCTCTCGC 60
Db 382 GAGAAGCTGAATGGAATTCACGAAGAGCACAATATACAACATATGTGATTCTTCGA 441
QY 61 CTTTGGAGAAACACATATTAAGAACAGTAGTCCCTTCTCTATAAAGAAAGCTAGTGTCT 120
Db 442 CTTTGGAGAAACACAGTACAAGAGCAGTAGTCCCTTCTCTACAGAGAGAGCTGGTGTCT 501
QY 121 GTAAACAAAGCTACCTCTACACAGACTATCAGGCTGGAGACGACAGATGTTTTCAGG 180
Db 502 GTGAAGACAAAATACCACATACCATGACTATCAGGATGGAGACACAGCGTGTTCAGG 561
QY 181 GAACCCCTTTGTGTGGTTCAGTCACTACACCGCTGTCAAGGACTTCGTGATTGTC 240
Db 562 GAGCCCTTTGTGTGGTTCCTTCCCTTACTGCTGTCAAGGACTTCGTGATTGTC 621
QY 241 CCCTGCACACACCCTCGTGAACATCCGTTAGAGAGATTGATGAGCTGGCTGTGATCTAC 300
Db 622 CCCTTGACACAACTCCGAGACCTCGTTAAAGAGATAGATGAGCTGGTGTGATCTAC 681
QY 301 ACAGATGTGAACGTGCTCGAATGCAGAGAAATTCATTTTATGGGTGACTTCAATGCT 360
Db 682 ACGGATGTGAGAAAGCCAGTGGAAAGACAGAGAAATTCATCTTCATGGGTGATTTCAACGCC 741
QY 361 GGCTGACGCTACGTCCTCCCAAGAGAGGCTGGAAGACATCCGCTGAGGACGGACCCCAAG 420
Db 742 GGCTGTAGCTATGTCCCAAGAGAGGCTGGCAGAACATTCGTTTGAGGACGGACCCCAAG 801
QY 421 TTCGTTTGGCTGATCGGGGACCAAGAGGACACACCGGTCAAGAGAGACACAAACTCGGCC 480
Db 802 TTTGTTTGGCTGATTTGGGACCAAGAGGACACTACGGTCAAGAGAGTACACGCTGTGCC 861
QY 481 TATGACAGGATCGTGTCTTAGAGGACAAATATATTGTCAACTCTCTGGTGTCTCTCAATCAAC 540
Db 862 TATGACAGGATTTGTGCTTTGTGGACAAAGAGATAGTCAACTCCGTGGTTCCTCCGTTCCAGT 921
QY 541 CTGCTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAAATCGAAGGCCCTGG 592
Db 922 GGCGTCTTTGACTTTTCAGAAAGCTTATGACTTGTCTGAGGAGGAGGCCCTGG 973
```

Search completed: December 9, 2002, 11:30:44

Job time : 1260.59 secs





Zhang H, Ax RL, Bellin ME;  
WPI; 2002-171803/22.  
P-PSDB; AAE18281.

Novel amino acid and polynucleotide sequence for fertility associated antigen useful for increasing stability of plasma membrane, acrosome and other portions of sperm cell, and increasing fertility of mammals

Claim 1; Fig 6; 37pp; English.

The invention relates to bovine fertility associated antigen (FAA) and its nucleic acid sequence. FAA is useful for increasing the stability of the plasma membrane plus acrosome of a sperm cell and/or other portions of a sperm cell, in particular mammalian sperm cell such as a sperm cell from buffalo, cow, horse, mice, pig, sheep, human, avian sperm cell including turkey or chicken sperm cell. FAA is also useful for increasing the fertility of sperm from a mammal including buffalo, cow, horse, mice, pig, sheep or human, or avian (chicken or turkey). FAA is added to the suspension of sperm cells prior to preservation or prior to administration of the sperm cells to a mammal in artificial insemination. Alternatively, fertility of sperm from a mammal is increased by placing FAA into a female's reproductive tract prior to deposition of the sperm cells into the female's reproductive tract by copulation or artificial insemination. The fertility of the male mammal is increased by injecting FAA into the male mammal's reproductive tract. The present sequence is bovine fertility associated antigen (FAA) partial cDNA.

Sequence 592 BP; 170 A; 143 C; 142 G; 137 T; 0 other;

Query Match	100.0%;	Score 592;	DB 24;	Length 592;
Best Local Similarity	100.0%;	Pred. No. 6e-178;		
Matches 592;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	GAGAAAGCTAAACCGAAATTC	CAAGAAAGGCATACATACAACTATG	TGATTAGCTCTCGC	60
Db 1	GAGAAAGCTAAACCGAAATTC	CAAGAAAGGCATACATACAACTATG	TGATTAGCTCTCGC	60
QY 61	CTTGGGAAGAACACATATAAGNACAGTATG	CTTCTTCTATAAGAAAGCTAGTGCT	120	
Db 61	CTTGGGAAGAACACATATAAGNACAGTATG	CTTCTTCTATAAGAAAGCTAGTGCT	120	
QY 121	GTAACAAAGAGTACCTCTACACAGACTAT	CAGGCTGGAGCGAGATGTGTTTCCAGG	180	
Db 121	GTAACAAAGAGTACCTCTACACAGACTAT	CAGGCTGGAGCGAGATGTGTTTCCAGG	180	
QY 181	GAACCTTTTGTTGGTTCAGTACACCTTAC	ACCGCTGTCAAGGACTTCGTTGATGTC	240	
Db 181	GAACCTTTTGTTGGTTCAGTACACCTTAC	ACCGCTGTCAAGGACTTCGTTGATGTC	240	
QY 241	CCCTGACACACACCCCTGAGACATCCGTT	AGAGATGATGAGCTGGCTGATGCTAC	300	
Db 241	CCCTGACACACACCCCTGAGACATCCGTT	AGAGATGATGAGCTGGCTGATGCTAC	300	
QY 301	ACAGATGTGAAGCTGCGTGGAAATTCAG	AGAAATTCATTTTCATGGGTGACTTCAATGCT	360	
Db 301	ACAGATGTGAAGCTGCGTGGAAATTCAG	AGAAATTCATTTTCATGGGTGACTTCAATGCT	360	
QY 361	GGCTGACGTACGTCTCCCAAGAGGCGCT	GGAAAGGCATCCGCCTGAGGACGACCCCAAG	420	
Db 361	GGCTGACGTACGTCTCCCAAGAGGCGCT	GGAAAGGCATCCGCCTGAGGACGACCCCAAG	420	
QY 421	TTGCTTTTGCTGATCGGGACCAAGAGCAC	CAACCGGTCAAGAGAGCACAACCTGCGCC	480	
Db 421	TTGCTTTTGCTGATCGGGACCAAGAGCAC	CAACCGGTCAAGAGAGCACAACCTGCGCC	480	
QY 481	TATGACAGGATCGTGTCTAGAGACAAAAT	TGTCAACTCTGGTGGTCTCAATCAAAAC	540	
Db 481	TATGACAGGATCGTGTCTAGAGACAAAAT	TGTCAACTCTGGTGGTCTCAATCAAAAC	540	
QY 541	CTCGTCTTTGATTTCCAGAAAGCTTAC	AGGTTGTCTGAATCGAAGGCCCTGG	592	

Db	541	CTCGCTTTTGATTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTCG	592
RESULT 2			
AAD29089			
ID	AAD29089	standard; DNA; 671 BP.	
XX	AC		
XX	AAD29089;		
XX			
DT	16-MAY-2002	(first entry)	
XX			
DE	Bovine 22kDa recombinant FAA (rFAA)	coding sequence.	
XX			
KW	Bovine; fertility associated antigen; FAA; sperm fertility; acrosome;		
KW	artificial insemination; plasma membrane; sperm cell; anti-infertility;		
KW	reproductive tract; ds.		
XX			
OS	Bos sp.		
PN	WO200206444-A2.		
XX			
PD	24-JAN-2002.		
XX			
PF	16-JUL-2001; 2001WO-US22097.		
XX			
PR	14-JUL-2000; 2000US-218140P.		
XX			
PA	(ARIZ-) ARIZONA BOARD OF REGENTS.		
XX			
PI	Zhang H, Ax RL, Bellin ME;		
XX			
DR	WPI; 2002-171803/22.		
XX			
PT	Novel amino acid and polynucleotide sequence for fertility associated		
PT	antigen useful for increasing stability of plasma membrane, acrosome		
PT	and other portions of sperm cell, and increasing fertility of mammals		
PT			
XX			
PS	Claim 1; Fig 5; 37pp; English.		
XX			
CC	The invention relates to bovine fertility associated antigen (FAA) and		
CC	its nucleic acid sequence. FAA is useful for increasing the stability of		
CC	the plasma membrane plus acrosome of a sperm cell and/or other portions		
CC	of a sperm cell, in particular mammalian sperm cell such as a sperm cell		
CC	from buffalo, cow, horse, mice, pig, sheep, human, avian sperm cell		
CC	including turkey or chicken sperm cell. FAA is also useful for increasing		
CC	the fertility of sperm from a mammal including buffalo, cow, horse, mice,		
CC	pig, sheep or human, or avian (chicken or turkey). FAA is added to the		
CC	suspension of sperm cells prior to preservation or prior to		
CC	administration of the sperm cells to a mammal in artificial insemination.		
CC	Alternatively, fertility of sperm from a mammal is increased by placing		
CC	FAA into a female's reproductive tract prior to deposition of the sperm		
CC	cells into the female's reproductive tract by copulation or artificial		
CC	insemination. The fertility of the male mammal is increased by injecting		
CC	FAA into the male mammal's reproductive tract. The present sequence is		
CC	bovine 22kDa recombinant FAA (rFAA) coding sequence.		
XX			
SQ	Sequence 671 BP; 195 A; 166 C; 154 G; 156 T; 0 other;		
	Query Match	100.0%; Score 592; DB 24; Length 671;	
	Best Local Similarity	100.0%; Pred. No. 6.5e-178;	
	Matches 592; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1	GAGAGCTAAACGGAAATTCAAGAAAAAGGCATACATACAACTATGTGATTAGCTTCGC	60
Db	27	GAGAGCTAAACGGAAATTCAAGAAAAAGGCATACATACAACTATGTGATTAGCTTCGC	86
Qy	61	CTTGAGAGAACACATATAAGAACATATGCCTTTCTCTATATAAGAAAGCTAGTGTCT	120
Db	87	CTTGAGAGAACACATATAAGAACATATGCCTTTCTCTATATAAGAAAGCTAGTGTCT	146
Qy	121	GTAAGAAACAAAGCTACCTCTACACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG	180

```
DB 147 GTAAACAAAGCTACCTCTACCAAGCTATCAGCTGAGACCAAGATGTGTTTCAGG 206
QY 181 GAACCTTTGTGCTGTGCTTCCAGTACCCCTTACACCGCTTCAAGAGCTTCGATGTC 240
DB 207 GAACCTTTGTGCTGTGCTTCCAGTACCCCTTACACCGCTTCAAGAGCTTCGATGTC 266
QY 241 CCCCTGACACCAACCCCTGAGACATCCGTTAGAGAGATGATGAGTGGCTGATGTCAC 300
DB 267 CCCCTGACACCAACCCCTGAGACATCCGTTAGAGAGATGATGAGTGGCTGATGTCAC 326
QY 301 ACAGATGTGAACGCTGCTGGAATGACAGAAATTTTCATGGGTGATCTTCAATGCT 360
DB 327 ACAGATGTGAACGCTGCTGGAATGACAGAAATTTTCATGGGTGATCTTCAATGCT 386
QY 361 GGTGACAGCTACCTCCCAAGAGAGCTGGAAGACATCCGCTGAGAGACGACCCCAAG 420
DB 387 GGTGACAGCTACCTCCCAAGAGAGCTGGAAGACATCCGCTGAGAGACGACCCCAAG 446
QY 421 TTGCTTTGGCTGATCGGGGACCAAGAGACACACGCTCAAGAAAGACACAACTGGGCC 480
DB 447 TTGCTTTGGCTGATCGGGGACCAAGAGACACACGCTCAAGAAAGACACAACTGGGCC 506
QY 481 TATGACAGAGATCGTGTCTTAGAGAGACAAATATTTGTCAACTCTGTGTCTCTCAATCAAC 540
DB 507 TATGACAGAGATCGTGTCTTAGAGAGACAAATATTTGTCAACTCTGTGTCTCTCAATCAAC 566
QY 541 CTGCTCTTTGATTTCCAGAAAGCTTACAGGTGTCTGAATTCGAAGGCCCTTGG 592
DB 567 CTGCTCTTTGATTTCCAGAAAGCTTACAGGTGTCTGAATTCGAAGGCCCTTGG 618
```

## RESULT 3

AAS17900

ID AAS17900 standard; cDNA, 578 BP.

XX AAS17900;

DT 26-MAR-2002 (first entry)

DE DNA encoding bovine fertility-associated antigen (FAA).

KW fertility-associated antigen; FAA; antifertility; sperm cell acrosome;  
KW reproductive fitness; single nucleotide polymorphism; SNP; bovine; cow;  
KW ss.

XX Bos taurus.

XX Key Location/Qualifiers

XX CDS 1..578

XX /tag= a

XX /product= "FAA"

XX /note= "Fertility-associated antigen"

XX /partial

XX /note= "No start or stop codon given"

XX /transl\_except= (pos:289..300, aa:A)

XX /note= This codon has an apparent 9 nucleotide

XX insertion"

XX /transl\_except= (pos:325..326, aa:R)

XX /note= "This codon has an apparent 1 nucleotide

XX deletion which alters the reading frame"

XX /transl\_except= (pos:576..578, aa:RSEKAL)

XX MO200176529-A2.

XX 18-OCT-2001.

XX 09-APR-2001; 2001MO-US10802.

XX 07-APR-2000; 2000US-195225P.

XX (ZHAN/) ZHANG H M.

XX (AXRL/) AX R L.

```
XX Zhang HW, Ax RL;
PI MPI: 2002-066307/09.
XX P-PSDB; AAU11516.
DR Novel human or bovine fertility-associated antigen useful for
XX stabilising sperm cell acrosome and increasing fertility of a male
PS Claim 25; Fig 2A-E; 54pp; English.
XX
XX The invention describes a novel isolated human fertility-associated
CC antigen (FAA) and a bovine FAA polypeptide. Human or bovine FAA is useful
CC for increasing the stability of a sperm cell acrosome and for increasing
CC the fertility of a human male by administering FAA. Detecting single
CC nucleotide polymorphisms (SNPs) of the human or bovine FAA gene is useful
CC for assaying fertility by detecting the presence or absence of a FAA SNP
CC in the mammal and correlating the presence or absence with the fertility
CC of the mammal. The FAA SNPs are chosen from FAA SNPs 1-34 given in the
CC specification. FAA SNP is also useful for determining the reproductive
CC fitness of a human or bovine, by determining the nucleotide sequence of
CC the FAA SNP, quantifying the fertility of more than one mammal containing
CC the FAA SNP and correlating the frequency of the FAA SNP to the
CC reproductive fitness of the mammals. FAA improves the integrity of sperm
CC membranes and increases the capacitation of sperm derived from either
CC fertile or infertile humans. Early genetic identification of infertility
CC improves the prognosis for subsequent attempts at fertilisation and
CC facilitates early intervention to determine whether the individual may
CC benefit from fertility treatments to avoid costly and/or emotional
CC problems with attempted inseminations. This sequence encodes the bovine
CC fertility-associated antigen (FAA) described in the method of the
CC invention.
XX
SQ Sequence 578 BP; 167 A; 140 C; 138 G; 133 T; 0 other;
```

Query Match 91.0%; Score 539; DB 24; Length 578;

Best Local Similarity 98.3%; Pred. No. 4.4e-161;

Matches 569; Conservative 0; Mismatches 0; Indels 10; Gaps 2;

```
QY 1 GAGAACTAAACGGAATTCAGAAAGGCTATACACTATGTGATTAGCTTCCG 60
DB 1 GAGAACTAAACGGAATTCAGAAAGGCTATACACTATGTGATTAGCTTCCG 60
QY 61 CTGGAAGAAACATATTAAGACAGTATGCTTCTATTAAGAAAGCTAGTGTCT 120
DB 61 CTGGAAGAAACATATTAAGACAGTATGCTTCTATTAAGAAAGCTAGTGTCT 120
QY 121 GTAAACAAAGCTACCTCTACACGACTATCAGGCTGAGAGCAGAGATGTTTCCAG 180
DB 121 GTAAACAAAGCTACCTCTACACGACTATCAGGCTGAGAGCAGAGATGTTTCCAG 180
QY 181 GAACCTTTGTGCTGTGCTTCCAGTACCCCTTACACCGCTGCAAGAGCTTCGATGTC 240
DB 181 GAACCTTTGTGCTGTGCTTCCAGTACCCCTTACACCGCTGCAAGAGCTTCGATGTC 240
QY 241 CCCCTGACACCAACCCCTGAGACATCCGTTAGAGAGATGAGCTGCTGAGCT 291
DB 241 CCCCTGACACCAACCCCTGAGACATCCGTTAGAGAGATGAGCTGCTGAGCT 300
QY 292 GATGTCTACACAGATGTGAAGCTGCTGGAATGACAGAAATTTTCATTTGAGTGAC 351
DB 292 GATGTCTACACAGATGTGAAGCTGCTGGAATGACAGAAATTTTCATTTGAGTGAC 359
QY 301 GATGTCTACACAGATGTGAAGCTGCTGGAATGACAGAAATTTTCATTTGAGTGAC 359
DB 301 GATGTCTACACAGATGTGAAGCTGCTGGAATGACAGAAATTTTCATTTGAGTGAC 359
QY 352 TTCAATGCTGCTGACGATACGCTCCCAAGAGAGCTGGAAGAGACATCCGCTGAGAGC 411
DB 352 TTCAATGCTGCTGACGATACGCTCCCAAGAGAGCTGGAAGAGACATCCGCTGAGAGC 419
QY 412 GACCCCAAGTGGTTGGTGTGATCGGGGACCAAGAGAGACACGCTCAAGAGAGCACA 471
DB 412 GACCCCAAGTGGTTGGTGTGATCGGGGACCAAGAGAGACACGCTCAAGAGAGCACA 479
QY 472 AACTGCGCTATGACAGATCGTGTCTTAGAGACAAATATTTGTCAACTCTGTGTCTCT 531
DB 472 AACTGCGCTATGACAGATCGTGTCTTAGAGACAAATATTTGTCAACTCTGTGTCTCT 531
```



XX WPI; 2002-426119/45.  
 XX  
 PT Diagnosing and detecting the progression of liver cancer,  
 PT hepatocellular carcinoma or metastatic liver tumor in a patient,  
 PT involves detecting the level of expression of two or more genes in a  
 PT liver tissue sample  
 XX  
 PS Claim 1; SEQ ID NO 3393; 298pp; English.  
 XX  
 CC The invention relates to a novel method for diagnosing and detecting the  
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver  
 CC tumor in a patient, and differentiating metastatic liver cancer from  
 CC hepatocellular carcinoma in a patient, involving detecting the level of  
 CC expression of two or more genes represented in ABN93503-ABN97455 in a  
 CC tissue sample. The method of the invention has hepatotropic, and  
 CC cytotropic activity. The method is useful for diagnosing and detecting  
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic  
 CC liver carcinoma in a patient. The method is useful for identifying  
 CC expression profiles which serve as useful diagnostic markers as well as  
 CC markers that can be used to monitor disease states, disease progression,  
 CC drug toxicity, drug efficacy and drug metabolism.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 SQ Sequence 1023 BP; 312 A; 244 C; 240 G; 227 T; 0 other;  
 Query Match 80.5%; Score 476.8; DB 24; Length 1023;  
 Best Local Similarity 87.8%; Pred. No. 3.9e-141;  
 Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;  
 Oy 1 GAGAGCTAAAGGAAATTCAGAAAGGCGTAACTATCACTATGATTAGCTCTCGC 60  
 Db 241 GAGAGCTAAAGGAAATTCAGAAAGGCGTAACTATGATTAGCTCTCGC 300  
 Oy 61 CTGGAAGAAACATATTAAGAAACAGTATGCTTTCTATTAAGAAAGCTATGCTCT 120  
 Db 301 CTGGAAGAAACATATTAAGAAACAGTATGCTTTCTATTAAGAAAGCTATGCTCT 360  
 Oy 121 GTAAAGAAAGCTATGCTTTCTATTAAGAAAGCTATGCTTTCTATTAAGAAAGCT 180  
 Db 361 GTAAAGAAAGCTATGCTTTCTATTAAGAAAGCTATGCTTTCTATTAAGAAAGCT 420  
 Oy 181 GAAAGCTTTGAGTCTGTTCCAGTCAACCGCTGACGCGTCAAGGAGCTTCGATGTC 240  
 Db 421 GAGCGCTTTGAGTCTGTTCCAGTCAACCGCTGACGCGTCAAGGAGCTTCGATGTC 480  
 Oy 241 CCCTGACACACACCCCTGAGACATCCGTTAGAGATTGATGAGTGGCTGATGCTAC 300  
 Db 481 CCCTGACACACACCCCTGAGACATCCGTTAGAGATTGATGAGTGGCTGATGCTAC 540  
 Oy 301 ACAGATGTAAGAGCTGCTGATGACAGAGATTTCATTTTCATGGGTGATCTCATGCT 360  
 Db 541 ACAGATGTAAGAGCTGCTGATGACAGAGATTTCATTTTCATGGGTGATCTCATGCT 600  
 Oy 361 GGCTGACAGTACGCTCCCAAGAGAGGCTGAGAGGACATCCGCTGAGAGGAGGAGCCCAAG 420  
 Db 601 GGCTGACAGTACGCTCCCAAGAGAGGCTGAGAGGACATCCGCTGAGAGGAGGAGCCCAAG 660  
 Oy 421 TTGCTTTGGCTGATCGGGGACCAAGAGAGACACACGCTGAAGAGAGACCACTGCGCA 480  
 Db 661 TTGCTTTGGCTGATCGGGGACCAAGAGAGACACACGCTGAAGAGAGACCACTGCGCA 720  
 Oy 481 TATGACAGAGTCTGCTTAAAGAGCAAAATATTGCTCACTCTGCTGCTCTCATCAAC 540  
 Db 721 TATGACAGAGTCTGCTTAAAGAGCAAAATATTGCTCACTCTGCTGCTCTCATCAAC 780  
 Oy 541 CTGCTCTTATTTCCAGAAAGCTTACAGGTTGCTGAATCGAAGGCGCCCTGG 592  
 Db 781 AGTCTTTTGAATTCAGAAAGCTTACAGGTTGCTGAATCGAAGGAGGCGCCCTGG 832

RESULT 6  
 ID ABL61914  
 AC ABL61914;  
 DT 15-MAY-2002 (first entry)  
 XX  
 DE Colon adenocarcinoma related gene sequence SEQ ID NO:251.  
 XX  
 KM Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
 KM stomach; lung; prostate; pancreas; carcinoma; antitumour; carcinos;  
 KM cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
 KM gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200194629-A2.  
 PD 13-DEC-2001.  
 XX  
 PF 30-MAY-2001; 2001WO-US10638.  
 XX  
 PR 05-JUN-2000; 2000US-209473P.  
 PR 05-JUN-2000; 2000US-209531P.  
 PR 18-SEP-2000; 2000US-233133P.  
 PR 18-SEP-2000; 2000US-233617P.  
 PR 20-SEP-2000; 2000US-234009P.  
 PR 20-SEP-2000; 2000US-234034P.  
 PR 20-SEP-2000; 2000US-234052P.  
 PR 22-SEP-2000; 2000US-234509P.  
 PR 22-SEP-2000; 2000US-234567P.  
 PR 25-SEP-2000; 2000US-234923P.  
 PR 25-SEP-2000; 2000US-234924P.  
 PR 25-SEP-2000; 2000US-235077P.  
 PR 25-SEP-2000; 2000US-235082P.  
 PR 25-SEP-2000; 2000US-235134P.  
 PR 25-SEP-2000; 2000US-235280P.  
 PR 26-SEP-2000; 2000US-235637P.  
 PR 26-SEP-2000; 2000US-235638P.  
 PR 27-SEP-2000; 2000US-235711P.  
 PR 27-SEP-2000; 2000US-235720P.  
 PR 27-SEP-2000; 2000US-235840P.  
 PR 27-SEP-2000; 2000US-235863P.  
 PR 28-SEP-2000; 2000US-236032P.  
 PR 28-SEP-2000; 2000US-236033P.  
 PR 28-SEP-2000; 2000US-236034P.  
 PR 28-SEP-2000; 2000US-236109P.  
 PR 28-SEP-2000; 2000US-236111P.  
 PR 29-SEP-2000; 2000US-236642P.  
 PR 29-SEP-2000; 2000US-236891P.  
 PR 02-OCT-2000; 2000US-237172P.  
 PR 02-OCT-2000; 2000US-237173P.  
 PR 02-OCT-2000; 2000US-237278P.  
 PR 02-OCT-2000; 2000US-237294P.  
 PR 02-OCT-2000; 2000US-237295P.  
 PR 02-OCT-2000; 2000US-237316P.  
 PR 03-OCT-2000; 2000US-237598P.  
 PR 03-OCT-2000; 2000US-237604P.  
 PR 03-OCT-2000; 2000US-237606P.  
 PR 03-OCT-2000; 2000US-237608P.  
 PR 01-NOV-2000; 2000US-244867P.  
 PR 01-NOV-2000; 2000US-245084P.  
 XX  
 PA (AVAL-) AVALON PHARM.  
 XX  
 PI Young PE, Augustus M, Carter KC, Edner R, Endress G, Horrigan S;  
 PI Soppet DR, Weaver Z;  
 XX  
 DR WPI; 2002-188264/24.  
 XX

PT Screening for anti-neoplastic agent involves exposing cells to a  
PT chemical agent to be tested for anti-neoplastic activity, and  
XX determining a change in expression of a gene of a signature gene set  
XX Claim 1; SEQ ID 251; 44pp; English.  
XX  
XX The present invention describes a method (M1) for screening for an  
CC anti-neoplastic agent. The method involves exposing cells to a chemical  
CC agent to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
CC to ABL70110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
CC activity and can be used in gene therapy. M1 can be used for screening  
CC an anti-neoplastic agent, and can be used for producing a product which  
CC is the data collected with respect to the anti-neoplastic agent as a  
CC result of M1, and the data is sufficient to convey the chemical  
CC structure and/or properties of the agent. M1 can be used in the  
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
CC carcinoma, papillary carcinoma and Wilm's tumour.  
XX  
SQ Sequence 1023 BP; 312 A; 244 C; 240 G; 227 T; 0 other;

Query Match 80.5%; Score 476.8; DB 24; Length 1023;  
Best Local Similarity 87.8%; Pred. No. 3.8e-141;  
Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;  
QY 1 GAGAGCTAAACGGAAATCAAGAAAGGCGATACATCAACTATGCTGATTAGCTCTCGC 60  
Db 241 GAGAGCTGAACAGAAATCAAGAGAGGCGATACATCAACTATGCTGATTAGCTCTCGC 300  
QY 61 CTGGAAGAAACACATATAAAGAACAGATGCTCTTCTTCTATAAGAAAGCTAGTGCT 120  
Db 301 CTGGAAGAAACACATATAAAGAACATATGCTCTTCTTCTCAAGAAAGCTGCTGCT 360  
QY 121 GTAAACAAAGCTACCTTACACAGCTATCAGGCTGAGAGCGAGATGTTTTCAGG 180  
Db 361 GTGAAGAGGAGTTACTACATGACTATCAGGATGAGAGCGAGATGTTTTCAGG 420  
QY 181 GAACCTTTGTGGTCTGGTTCAGTCACTCCCTACACCGCTGTCAGGACTTCGTGATTGTC 240  
Db 421 GAGCCCTTTGTGGTCTGGTTCATCTCCACACTGCTGTCAGAGACTTCGTGATTATC 480  
QY 241 CCCCTGCACACACCCCTGAGACATCCGTTAGAGAGATGATGAGCTGGCTGATGCTTAC 300  
Db 481 CCCCTGCACACACCCCTGAGACATCCGTTAAGGAGATCGATGATGTTGGTGGCTTAC 540  
QY 301 ACAGATGTAAACGTCGCTGGAATGCAGAGAAATTCATTTTCATGGGTGACTTCAATGCT 360  
Db 541 ACGGACGTGAACACCCGCTGGAAGCGGAGAAATTCATTTTCATGGGTGACTTCAATGCC 600  
QY 361 GGCTGCAGCTACGCTCCCAAGAGCCCTGGAAGGACATCCGCTGAGGACGACCCCAAG 420  
Db 601 GGCTGCAGCTACGCTCCCAAGAGCCCTGGAAGGACATCCGCTGAGGACTGACCCCAAG 660  
QY 421 TTCGTTGGCTGATCGGGGACCAAGAGACACACGCTGGAAGAGAGACAAACTGCGCC 480  
Db 661 TTTGTTGGCTGATCGGGGACCAAGAGACACACGCTGGAAGAGAGACCAACTGCGCA 720  
QY 481 TATGACAGATCGCTTTAGAGGACAAATATTCCTGCTGCTGCTGCTCAATCAAC 540  
Db 721 TATGACAGATCGCTTTAGAGGACAAATATTCCTGCTGCTGCTGCTCAATCAAC 780  
QY 541 CTCGCTTTGATTTTCAGAAAGCTTACAGGTTGCTGTAATCGAAGGCCCTGG 592  
Db 781 AGTGTTTGTGACTCCAGAAAGCTTACAGCTGACTGAAGAGAGGCCCTGG 832

RESULT 7  
ABL64458

ID ABL64458 standard; DNA; 1023 BP.  
XX  
AC ABL64458;  
XX  
DT 15-MAY-2002 (first entry)  
XX  
XX Stomach cancer related gene sequence SEQ ID NO:2795.  
DE Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US10838.  
XX  
PR 05-JUN-2000; 2000US-209473P.  
PR 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 18-SEP-2000; 2000US-233617P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 20-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 22-SEP-2000; 2000US-234567P.  
PR 25-SEP-2000; 2000US-234923P.  
PR 25-SEP-2000; 2000US-234924P.  
PR 25-SEP-2000; 2000US-235077P.  
PR 25-SEP-2000; 2000US-235082P.  
PR 25-SEP-2000; 2000US-235134P.  
PR 25-SEP-2000; 2000US-235280P.  
PR 26-SEP-2000; 2000US-235637P.  
PR 26-SEP-2000; 2000US-235638P.  
PR 27-SEP-2000; 2000US-235711P.  
PR 27-SEP-2000; 2000US-235720P.  
PR 27-SEP-2000; 2000US-235840P.  
PR 27-SEP-2000; 2000US-235863P.  
PR 28-SEP-2000; 2000US-236028P.  
PR 28-SEP-2000; 2000US-236032P.  
PR 28-SEP-2000; 2000US-236033P.  
PR 28-SEP-2000; 2000US-236034P.  
PR 28-SEP-2000; 2000US-236109P.  
PR 28-SEP-2000; 2000US-236111P.  
PR 29-SEP-2000; 2000US-236842P.  
PR 29-SEP-2000; 2000US-236891P.  
PR 02-OCT-2000; 2000US-237172P.  
PR 02-OCT-2000; 2000US-237173P.  
PR 02-OCT-2000; 2000US-237278P.  
PR 02-OCT-2000; 2000US-237294P.  
PR 02-OCT-2000; 2000US-237295P.  
PR 02-OCT-2000; 2000US-237316P.  
PR 03-OCT-2000; 2000US-237425P.  
PR 03-OCT-2000; 2000US-237598P.  
PR 03-OCT-2000; 2000US-237604P.  
PR 03-OCT-2000; 2000US-237606P.  
PR 03-OCT-2000; 2000US-237608P.  
PR 01-NOV-2000; 2000US-244867P.  
PR 01-NOV-2000; 2000US-245084P.  
XX  
PA (AVAL-) AVALON PHARM.  
XX  
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
PI Soppet DR, Weaver Z;  
XX  
DR WPI; 2002-188264/24.  
XX  
PT Screening for anti-neoplastic agent involves exposing cells to a  
PT chemical agent to be tested for anti-neoplastic activity, and

PT determining a change in expression of a gene of a signature gene set -  
XX  
PS Claim 1; SEQ ID 2795; 4app; English.  
XX  
CC The present invention describes a method (M1) for screening for an  
CC anti-neoplastic agent. The method involves exposing cells to a chemical  
CC agent to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
CC to ABL7010), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
CC activity and can be used in gene therapy. M1 can be used for screening  
CC an anti-neoplastic agent, and can be used for producing a product which  
CC is the data collected with respect to the anti-neoplastic agent as a  
CC result of M1, and the data is sufficient to convey the chemical  
CC structure and/or properties of the agent. M1 can be used in the  
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
CC carcinoma, papillary carcinoma and Wilms' tumour.  
XX  
SQ Sequence 1023 BP; 312 A; 244 C; 240 G; 227 T; 0 other;  
  
Query Match 80.5%; Score 476.8; DB 24; Length 1023;  
Best Local Similarity 87.8%; Pred. No. 3.8e-141;  
Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;  
  
QY 1 GAGACCTAAACCGAATTCAAGAAAGGCAATACATACATCTGTGATTAGTCTCCG 60  
DB 241 GAGAACTGACAGAAATTTCAAGAGAGGCAATACATACATCTGTGATTAGTCTCCG 300  
QY 61 CTGGAAGAAACATATTAAGAAACAGTATGCTCTTCTATTAAGAAAGCTAGTCT 120  
DB 301 CTGGAAGAAACATATTAAGAAACATATGCTCTTCTATTAAGAAAGCTAGTCT 360  
QY 121 GTAAAAAAGGCTACTCTTACACGACTATCAGCTGAGAGCGAGATGTGTTCCAGG 180  
DB 361 GTGAAGAGAGTATATCACTACATGACTATCAGATGAGACCGAGATGTGTTCCAGG 420  
QY 181 GAACCTTTGT 240  
DB 421 GAACCTTTGT 480  
QY 241 CCCCTGACACCAACCCCTGAGATCCGTTAGAGATGATGAGTGTGTGTGTGTGTGT 300  
DB 481 CCCCTGACACCAACCCCTGAGATCCGTTAGAGATGATGAGTGTGTGTGTGTGTGT 540  
QY 301 ACAGATGTGAAGCGT 360  
DB 541 ACAGATGTGAAGCGT 600  
QY 361 GGTGAGAGTACGTCCTCCCAAGAGAGGCTGTGAGAGACATCCGCTTGAAGAGAGG 420  
DB 601 GGTGAGAGTACGTCCTCCCAAGAGAGGCTGTGAGAGACATCCGCTTGAAGAGAGG 660  
QY 421 TTGCTGTGCTGATGCGGGGACCAAGAGACACACGCTCAAGAGAGCAAACTGCGCC 480  
DB 661 TTGCTGTGCTGATGCGGGGACCAAGAGACACACGCTCAAGAGAGCAAACTGCGCA 720  
QY 481 TATGACAGATCGT 540  
DB 721 TATGACAGATCGT 780  
QY 541 CTCGTCTTTGATTTCCAAAGCTTACAGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 592  
DB 781 AGT 832  
  
RESULT 8  
ABL64961  
ID ABL64961 standard; DNA; 1023 BP.  
XX

AC ABL64961;  
XX  
DT 15-MAY-2002 (First entry)  
XX  
DE Lung cancer related gene sequence SEQ ID NO:3298.  
XX  
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;  
KW gene; dr.  
XX  
OS Homo sapiens.  
XX  
PN W0200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US10838.  
XX  
PR 05-JUN-2000; 2000US-209473P.  
PR 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 18-SEP-2000; 2000US-233137P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 20-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 22-SEP-2000; 2000US-234567P.  
PR 25-SEP-2000; 2000US-234923P.  
PR 25-SEP-2000; 2000US-234924P.  
PR 25-SEP-2000; 2000US-235077P.  
PR 25-SEP-2000; 2000US-235082P.  
PR 25-SEP-2000; 2000US-235134P.  
PR 25-SEP-2000; 2000US-235280P.  
PR 26-SEP-2000; 2000US-235637P.  
PR 26-SEP-2000; 2000US-235638P.  
PR 27-SEP-2000; 2000US-235711P.  
PR 27-SEP-2000; 2000US-235720P.  
PR 27-SEP-2000; 2000US-235840P.  
PR 27-SEP-2000; 2000US-235863P.  
PR 28-SEP-2000; 2000US-236028P.  
PR 28-SEP-2000; 2000US-236032P.  
PR 28-SEP-2000; 2000US-236033P.  
PR 28-SEP-2000; 2000US-236034P.  
PR 28-SEP-2000; 2000US-236109P.  
PR 28-SEP-2000; 2000US-236111P.  
PR 29-SEP-2000; 2000US-236812P.  
PR 29-SEP-2000; 2000US-236891P.  
PR 02-OCT-2000; 2000US-237172P.  
PR 02-OCT-2000; 2000US-237173P.  
PR 02-OCT-2000; 2000US-237278P.  
PR 02-OCT-2000; 2000US-237294P.  
PR 02-OCT-2000; 2000US-237295P.  
PR 02-OCT-2000; 2000US-237316P.  
PR 03-OCT-2000; 2000US-237425P.  
PR 03-OCT-2000; 2000US-237598P.  
PR 03-OCT-2000; 2000US-237604P.  
PR 03-OCT-2000; 2000US-237606P.  
PR 03-OCT-2000; 2000US-237608P.  
PR 01-NOV-2000; 2000US-244867P.  
PR 01-NOV-2000; 2000US-245084P.  
XX  
PA (AVAL-) AVALON PHARM.  
XX  
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S,  
PI Soppet DR, Weaver Z;  
XX  
XX WPI; 2002-188264/24.  
XX  
PT Screening for anti-neoplastic agent involves exposing cells to a  
PT chemical agent to be tested for anti-neoplastic activity, and  
PT determining a change in expression of a gene of a signature gene set -  
XX

PS Claim 1; SEQ ID 3298; 44pp; English.

XX The present invention describes a method (M1) for screening for an

CC anti-neoplastic agent. The method involves exposing cells to a chemical

CC agent to be tested for anti-neoplastic activity, determining a change in

CC expression of at least one gene (I) of a signature gene set, where (I)

CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664

CC to ABL70110), or is at least 95% identical to (S), where a change in

CC expression is indicative of anti-neoplastic activity. (I) has cytostatic

CC activity and can be used in gene therapy. M1 can be used for screening

CC an anti-neoplastic agent, and can be used for producing a product which

CC is the data collected with respect to the anti-neoplastic agent as a

CC result of M1, and the data is sufficient to convey the chemical

CC structure and/or properties of the agent. M1 can be used in the

CC treatment of cancer such as colon, breast, stomach, lung, thyroid,

CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,

CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,

CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine

CC carcinoma, papillary carcinoma and Wilms' tumour.

XX

SQ Sequence 1023 BP; 312 A; 244 C; 240 G; 227 T; 0 other;

Query Match 80.5%; Score 476.8; DB 24; Length 1023;

Best Local Similarity 87.8%; Pred. No. 3.8e-141;

Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 1 GAGAAGCTAAACGGAAATTCAGAAAGGCGCATACATCAACTATGTGATTAGCTCTCGC 60

DB 241 GAGAAGCTGAACAGAAATTCAGGAGAGGCATACGTACACTATGTGATTAGCTCTCGG 300

QY 61 CTTGGAAGAACACATATAAGAACAGATATGCTTTCTCTATAAGAAAGCTAGTGTCT 120

DB 301 CTTGGAAGAAACACATATAAGAAACATATGCTTTCTCTATAAGAAAGCTAGTGTCT 360

QY 121 GTAAACAAAGCTACTCTTACACACATATCAGGCTGAGAGCGAGATGTGTTTCCAGG 180

DB 361 GTGAAGAGAGTTATCATCTACATGACTATCAGGATGAGAGCGAGATGTGTTTCCAGG 420

QY 181 GAACCCCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 240

DB 421 GAGCCCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 480

QY 241 CCCCTGCACACACCCCTGCAGACATCCGTTAGAGAGATGATGAGCTGGCTGATGCTTAC 300

DB 481 CCCCTGCACACACCCCTGCAGACATCCGTTAAGGAGATCGATGATGTTGGTTGAGGCTTAC 540

QY 301 ACAGATGTGAACCGTCTGCTGGAATGCAGAGAAATTTCAATTTTCATGGGTGACTTCAATGCT 360

DB 541 ACGGACGTGAACACCGCTGGAGCGGAGAAATTTCAATTTTCATGGGTGACTTCAATGCC 600

QY 361 GGCTGCAGCTACGTCCTCCCAAGAGCGCTGGAGGACATCCGCTTGAGACCGGACCCCAAG 420

DB 601 GGCTGCAGCTACGTCCTCCCAAGAGCGCTGGAGGACATCCGCTTGAGGACTGACCCCAAG 660

QY 421 TTCGTTGGCTGATCGGGGCAAGAGACACCCAGGTCAGAGAGACACAACTCGGCC 480

DB 661 TTTGTTGGCTGATCGGGGCAAGAGACACCCAGGTCAGAGAGACACAACTCGGTCGA 720

QY 481 TATGACGATGCTGCTTAGAGGACAAATATTTGTCACCTCTGGTGGTCTCAATCAAAAC 540

DB 721 TATGACGATGCTGCTTAGAGGACAAATATTTGTCACCTCTGGTGGTCTCAATCAAAAC 780

QY 541 CTCGCTTTGATTTCAGAAAGCTTACAGGTTGCTGTAATCGAAGGCCCTCG 592

DB 781 AGTGTGTTTGTGCTTCCAGAAAGCTTACAAAGCTGACTGAAGAGGAGGCCCTCG 832

RESULT 9

AAT74076

ID AAT74076 standard; cDNA; 1079 BP.

XX

AC AAT74076;

XX

DT 29-JAN-1998 (first entry)

XX Human LS-DNase nucleotide sequence.

DE

XX DNase; actin; DNA viscoelasticity; systemic lupus erythematosus;

XX cystic fibrosis; meningitis; pulmonary disease; gene therapy;

KW recombinant preparation; DNA-hydrolytic activity; ss.

KW

XX Homo sapiens.

OS

XX Key Location/Qualifiers

PH 71..988

FT CDS

FT sig\_peptide /tag= a

FT mat\_peptide /tag= b

FT /tag= c

FT /product= LS-DNase

XX

PN WO9728266-A1.

XX

XX 07-AUG-1997.

XX

XX 03-FEB-1997; 97WO-US01506.

PF

XX 05-FEB-1996; 96US-0597078.

PR

XX (GETH ) GENENTECH INC.

PA

XX Baker KP, Baron WF;

PI

XX WPI; 1997-402619/37.

DR

XX P-PSDB; AAW23542.

XX

XX New isolated LS-DNase which is resistant to inhibition by actin -

PT used for the treatment of e.g systemic lupus erythematosus and

PT pulmonary diseases and disorders such as cystic fibrosis

DR

XX Claim 1; Fig 1; 32pp; English.

PS

XX

CC This DNA sequence encodes a novel human LS-DNase (AAW23542), which has

CC DNA-hydrolytic activity but is resistant to inhibition by actin. It can

CC be used for reducing the viscoelasticity or viscous consistency of

CC DNA-containing material, as well as treatment of systemic lupus

CC erythematosus or pulmonary disease of disorders such as cystic fibrosis,

CC bronchitis, pneumonia, bronchiectasis, emphysema, asthma, tuberculosis

CC or fungal infections. It can also be used for adjunctive treatment of

CC abscesses or severe closed-space infections in conditions such as

CC empyema, meningitis, peritonitis, sinusitis, otitis, periodontitis,

CC pericarditis, pancreatitis, cholelithiasis, endocarditis and

CC septic arthritis, as well as in topical treatments in a variety of

CC inflammatory and infected lesions, e.g. of the skin and mucosal

CC membranes, surgical wounds, ulcerative lesions and burns. LS-DNase may

CC improve the efficacy of antibiotics used in the treatment of such

CC infections (e.g. gentamicin activity is markedly reduced by reversible

CC binding to intact DNA). It can also be used for the treatment of other

CC non-infected conditions in which there is an accumulation of cellular

CC debris that includes cellular DNA, such as pyelonephritis and

CC tubulo-interstitial kidney disease. It can also be used in vitro

CC diagnostic assays of a viscous material (e.g. sputum) from a patient.

CC The anti-LS-DNase antibodies can be used for detection and purification

CC of LS-DNase.

XX

SQ Sequence 1079 BP; 330 A; 260 C; 252 G; 237 T; 0 other;

Query Match 80.5%; Score 476.8; DB 18; Length 1079;

Best Local Similarity 87.8%; Pred. No. 3.9e-141;

Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 1 GAGAAGCTAAACGGAAATTCAGAAAGGCGCATACATCAACTATGTGATTAGCTCTCGC 60

DB 287 GAGAAGCTGAACAGAAATTCAGGAGAGGCGCATACATCAACTATGTGATTAGCTCTCGG 346



```

QY 61 CTTGGAAGAAACATATTAAGAAACAGTAGCCCTTTCTCTATTAAGAAAGACTAGTCT 120
DB 347 CTTGGAAGAAACATATTAAGAAACATATGCTTTCTCTACAGAAAGAAAGCTGCTCT 406
QY 121 GTAAACAAAGCTACCTCTACAGACTATACAGCTGGAGCGAGATGTGTCTTCCAGG 180
DB 407 GTGAAGAGAGATTATCACTACATGACTATCAGATGAGAGCGAGATGTGTCTTCCAGG 466
QY 181 GAACCTTTTGCTGTGCTGTTCCAGTCAACCCGCTGCAAGGACTTGTGATGTC 240
DB 467 GAGCCCTTTTGCTGTGCTGTTCCAGTCAACCCGCTGCAAGGACTTGTGATGTC 526
QY 241 CCCCTGACACACCCCTGAGACATCCGTTAGAGATGATGAGCTGCTGATGTCTAC 300
DB 527 CCCCTGACACACCCCTGAGACATCCGTTAGAGATGATGAGCTGCTGATGTCTAC 586
QY 301 ACAGATGTGAACGCTGCTGATGATGAGAGAAATTTCAATTTCAATGAGTCTCAATGCT 360
DB 587 ACAGATGTGAACGCTGCTGATGATGAGAGAAATTTCAATTTCAATGAGTCTCAATGCT 646
QY 361 GGCTGACGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 647 GGCTGACGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 706
QY 421 TTGCTTTGCTGATCGGAGACCAAGAGACACACGCTGCAAGAGAGACCAACTGCGCC 480
DB 707 TTGCTTTGCTGATCGGAGACCAAGAGACACACGCTGCAAGAGAGACCAACTGCGCA 766
QY 481 TATGACAGGATCGCTGCTTACAGACAAATTTGTCAACTGCTGCTGCTCAATCAAC 540
DB 767 TATGACAGGATCGCTGCTTACAGACAAATTTGTCAACTGCTGCTGCTCAATCAAC 826
QY 541 CTGCTCTTATTTCCAGAAAGCTTACAGGTTGCTGATGAGAGGCGCTGG 592
DB 827 AGTGTTTTTCCTTCCAGAAAGCTTACAGGTTGCTGATGAGAGGCGCTGG 878

RESULT 10
ABK84429
ID ABK84429 standard; cDNA; 1108 BP.
AC ABK84429;
XX
XX
DT 14-AUG-2002 (first entry)
XX
DE Human cDNA differentially expressed in granulocytic cells #1000.
XX
XX Human; ss; granulocytic cell; DNA chip; bacterial infection;
XX viral infection; parasitic infection; protozoal infection;
XX fungal infection; sterile inflammatory disease; psoriasis;
XX rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
XX cardiac reperfusion injury; renal reperfusion injury; ARDS;
XX adult respiratory distress syndrome; inflammatory bowel disease;
XX Crohn's disease; ulcerative colitis; peridontal disease;
XX granulocyte activation; chronic inflammation; allergy.
XX
XX Homo sapiens.
XX
XX WO200228999-A2.
XX
XX 11-APR-2002.
XX
XX 03-OCT-2001; 2001WO-US30821.
XX
XX 03-OCT-2000; 2000US-237189P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX
XX MPI, 2002-435328/46.
XX
XX Detecting granulocyte activation by detecting differential expression

```

```

PT of genes associated with granulocyte activation, which serves as
PT diagnostic markers that is useful for monitoring disease states and
PT drug toxicity -
XX
XX Claim 1; SEQ ID No 1000; 114bp; English.
XX
XX The invention relates to detecting (M1) granulocyte (GC) activation
XX (GCA), by detecting the level of expression of gene(s) (Gs) identified by
XX DNA chip analysis as given in the specification, and comparing
XX the expression level to an expression level in an unactivated
XX GC, where differential expression of Gs is indicative of GCA.
XX Also included are modulating (M2) GA by contacting Gs with an agent
XX that alters the expression of at least one gene in Gs; (2) screening
XX for an agent capable of modulating GCA or an inflammation (especially
XX chronic) in a tissue, an allergic response in a subject, exposure of a
XX subject to a pathogen or sterile inflammatory disease using the
XX gene expression profile; (3) detecting (M4) an inflammation (especially
XX chronic) in a tissue, an allergic response in a subject, exposure of a
XX subject to a pathogen or sterile inflammatory disease, by detecting the
XX level of expression in a sample of the tissue of gene(s) from Gs, where
XX the level of expression of the gene is indicative of inflammation;
XX (4) treating (M5) an inflammation (especially chronic) or in a tissue,
XX an allergic response in a subject, exposure of a subject to a pathogen
XX or sterile inflammatory disease, by contacting a tissue having
XX inflammation with an agent that modulates the expression of gene(s)
XX from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
XX modulating GA; M3 is useful for screening an agent capable of modulating
XX GCA preferably in an inflammation in a tissue; M4 is useful for
XX detecting an inflammation (especially chronic) in a tissue, an allergic
XX response in a subject, exposure of a subject to a pathogen or sterile
XX inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
XX glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
XX reperfusion injury, ARDS, adult respiratory distress syndrome,
XX inflammatory bowel disease, Crohn's disease, ulcerative colitis,
XX peridontal disease; also bacterial infection, viral infection,
XX parasitic infection, protozoal infection, fungal infection and M5 is
XX useful for treating one of the above conditions. The present
XX sequence represents a gene differentially expressed in granulocytes.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic
XX format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1108 BP; 329 A; 261 C; 271 G; 247 T; 0 other;
Query Match 80.3%; Score 475.2; DB 24; Length 1108;
Best Local Similarity 87.7%; Pred. No. 1,3e-140;
Matches 519; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
QY 1 GAGAGCTAAACGGAATTTCAAGAAAGCATATACATATGATGATGCTCCG 60
DB 309 GAGAGCTAAACGGAATTTCAAGAAAGCATATGATGATGCTCCG 368
QY 61 CTTGGAAGAAACATATTAAGAAACAGTAGCCCTTTCTCTATTAAGAAAGACTAGTCT 120
DB 369 CTTGGAAGAAACATATTAAGAAACATATGCTTTCTCTACAGAAAGAGCTGCTCT 428
QY 121 GTAAACAAAGCTACCTCTACAGACTATCAGCTGAGAGCCAGATGTGTCTCCAGG 180
DB 429 GTGAAGAGAGATTATCACTACATGACTATCAGATGAGAGCCAGATGTGTCTCCAGG 488
QY 181 GAACCTTTTGCTGTGCTGTTCCAGTCAACCCGCTGCAAGGACTTGTGATGTC 240
DB 489 GAGCCCTTTTGCTGTGCTGTTCCAGTCAACCCGCTGCAAGGACTTGTGATGTC 548
QY 241 CCCCTGACACACCCCTGAGACATCCGTTAGAGATGATGAGCTGCTGATGTCTAC 300
DB 549 CCCCTGACACACCCCTGAGACATCCGTTAGAGATGATGAGCTGCTGATGTCTAC 608
QY 301 ACAGATGTGAACGCTGCTGATGATGAGAGAAATTTCAATTTCAATGAGTCTCAATGCT 360
DB 609 ACAGATGTGAACGCTGCTGATGATGAGAGAAATTTCAATTTCAATGAGTCTCAATGCT 668

```

QY 361 GGCTGAGCTACGTCCTCCCAAGAGCCCTGGGAAGACATCCGCTTGGAGCGGACCCCAAG 420  
Db |||||  
QY 669 GGCTGAGCTACGTCCTCCCAAGAGCCCTGGGAAGACATCCGCTTGGAGCTGACCCCAAG 728  
Db |||||  
QY 421 TTCGTTTGGCTGATCGGGACCAAGAGGACACACCGGTCAAGAAAGAGCACAACCTGGCC 480  
Db |||||  
QY 481 TATGACAGGATCGTCTTAGAGGACAAATAATTGTCAACTCTGGTGTCTCAATCAAC 540  
Db |||||  
QY 789 TATGACAGGATGTGCTTAGAGGACAAATAATCGTGTCTGTTGTTCCCAAGTCAAC 848  
QY 541 CTCGCTTTGATTCCAGAAAGCTTACAGGTGTCGAAATCGAAGGCCCTGG 592  
Db |||||  
QY 849 AGTGTGTTTGATCTCCAGAAAGCTTACAGCTGACTGAAGAGGAGGCCCTGG 900  
Db |||||

## RESULT 11

AAT17733  
ID AAT17733 standard; cDNA; 1208 BP.  
XX AC AAT17733;  
XX 17-OCT-1996 (first entry)  
XX DE Deoxyribonuclease, DNase-gamma cDNA.  
XX KW Deoxyribonuclease; DNase-gamma; selective cleavage; linker;  
KW chromatin DNA; cell nucleus; 3'-hydroxy; 5'-phosphate; apoptosis;  
KW carcinogenesis; AIDS development; diagnosis; prevention;  
KW treatment; cancer; autoimmune disease; viral infections; ds.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
XX CDS 12..944  
XX FT /\*tag= a  
XX FT /note= "GAT codon comprising bases 837-839 has  
FT no corresponding amino acid in AAR94021"

W09607735-A1.

14-MAR-1996.

06-SEP-1995; 95WO-JP01775.

06-SEP-1994; 94JP-0239518.

(TANU/) TANUMA S.

Tanuma S;

WPI; 1996-171610/17.

P-PSDB; AAR94021.

XX DNase which selectively cuts linker region of chromatin DNA -  
XX useful in diagnosis, treatment and prevention of cancer, autoimmune  
XX diseases, viral infections, etc.

PS Claim 4; Pages 55-57; 75pp; Japanese.

XX The present sequence encodes the deoxyribonuclease, DNase-gamma,  
XX which is able to selectively cleave the linker part of chromatin  
XX DNA. DNase-gamma has a mol. wt. of 32000 by SDS-PAGE, or 30000 by  
XX gel filtration, an optimum pH of 5.6, is found in the cell nucleus,  
XX is not dependent on divalent cations, has a Zn(2+) inhibition  
XX IC(50) > 1mM and leaves 3'-OH and 5'-P termini after cleavage. The  
XX DNase can be used at a mol. level to resolve the control  
XX mechanisms of carcinogenesis, the autoimmune system and AIDS  
XX development. It can also be used in the diagnosis, prevention and  
XX treatment of cancer, autoimmune disease, viral infections, etc.,  
XX and as an apoptosis control agent.

SQ Sequence 1208 BP; 319 A; 304 C; 282 G; 303 T; 0 other;

Query Match 76.2%; Score 451.2; DB 17; Length 1208;  
Best Local Similarity 85.1%; Pred. No. 5.8e-133;  
Matches 504; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 1 GAGAAGCTAAACCGAAATTCAAGAAAGGCATACATACAACTATGTAGTCTCTCGC 60  
Db |||||  
QY 243 GAGAAGCTGAATGAAACTCAGGAAGACGACATACAACTACGTGATGTCTCGG 302  
Db |||||  
QY 61 CTTGGAAGAAACATATAAAGAACAGTATGCTCTTCTATATAAAGAAAGCTAGTGTCT 120  
Db |||||  
QY 303 CTTGGAAGAAACATATAAAGAACAGTATGCTCTTCTATAAAGAAAGCTGTGTCT 362  
QY 121 GTAAACAAAGCTACCTCTACACGACTATCAGGCTGGAGACGAGATGTGTTTCAGG 180  
Db |||||  
QY 363 GTGAAGGCAAAATACCTCTACCATGACTATCAGGATGGAGACACAGACGTGTTTCCAGG 422  
QY 181 GAACCCCTTTGTGCTGCTGCTCCAGTCACCTACACGCTCTCAAGGACTTCGTGATTGTC 240  
Db |||||  
QY 423 GAGCCCTTTGTGCTGCTGCTCCAGGCGCCCTTCACTGCTCCCAAGGACTTCGTGATTGTC 482  
QY 241 CCCTGCACACCAACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGTGCTAC 300  
Db |||||  
QY 483 CCCTGCACACCAACTCTGAAACCTCTGTTAAAGAGATAGATGAGCTGGCTGAGCTCTAC 542  
QY 301 ACAGATGTGAACGCTGCTCGAATGAGAGAAATTCATTTTCATGGTGACTTCAATGCT 360  
Db |||||  
QY 543 ACAGATGTGTAAGAGACGATGGAAGGACAGATTTTCATCTTTCATGGTGATTTCAATGCT 602  
QY 361 GGCTGACAGCTACCTCCCAAGAGGCTGGAAGGACATCCGCTGAGGAGGACCCCAAG 420  
Db |||||  
QY 603 GGCTGACAGCTACCTCCCAAGAGGCTGGAAGGACATCCGCTTGAAGAGAGAGAGAGAG 662  
QY 421 TTGCTTTGGCTGATCGGGACCAAGAGGACACCAACGCTCAAGAGACACAACTGCGCC 480  
Db |||||  
QY 663 TTTGTTGGCTGATTTGGGACCAAGAGGACACCAACGCTCAAGAGAGACACAGCTGTGCC 722  
QY 481 TATGACAGGATCGTGTGCTTAGAGGACAAATAATTGTCAACTCTCTGCTGTGCTCTCAATCAAC 540  
Db |||||  
QY 723 TATGACAGGATTCGCTTCGCGGACCAAGAGATAGTCAACTCTCTGCTGTGCTGCTCCAGT 782  
QY 541 CTGCTCTTTCATTTCCAGAAAGCTTACAGGTTGCTGCTCAATCGAAGGCCCTGG 592  
Db |||||  
QY 783 GCGCTCTTTCAGAAAGCTTATGAGTTGCTGTAAGAGAGAGGCCCTGG 834  
Db |||||

## RESULT 12

AAT74083  
ID AAT74083 standard; cDNA; 1124 BP.  
XX AC AAT74083;  
XX 29-JAN-1998 (first entry)  
XX DE Murine LS-DNase nucleotide sequence.  
XX KW DNase; actin; DNA viscoelasticity; systemic lupus erythematosus;  
KW cystic fibrosis; meningitis; pulmonary disease; gene therapy;  
KW recombinant preparation; DNA-hydrolytic activity; ss.  
XX OS Mus sp.  
XX FH Key Location/Qualifiers  
XX CDS 173..1105  
XX FT /\*tag= a  
XX FT sig\_peptide 173..247  
XX FT /\*tag= b  
XX FT /note= "putative"  
XX FT mat\_peptide 248..1105  
XX FT /\*tag= c  
XX FT /product= LS-DNase

PN W09128266-A1.  
 XX 07-AUG-1997.  
 PD  
 XX  
 PF 03-FEB-1997; 97MO-US01506.  
 XX  
 XX 05-FEB-1996; 96US-0597078.  
 PR  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Baker KP, Baron WF;  
 XX  
 XX WPI; 1997-402619/37.  
 DR  
 XX  
 XX  
 PT New isolated LS-DNase which is resistant to inhibition by actin -  
 PT used for the treatment of e.g systemic lupus erythematosus and  
 PT pulmonary diseases and disorders such as cystic fibrosis  
 PS  
 XX Example 6; Fig 3; 32pp; English.  
 XX

CC This DNA sequence encodes a variant LS-DNase, which has DNA-hydrolytic  
 CC activity but is resistant to inhibition by actin. A 649 base pair  
 CC fragment of the coding sequence of cloned human LS-DNase (AAT74076) was  
 CC used to screen a murine liver cDNA library and obtain this nucleotide.  
 CC Murine mature LS-DNase is 84% identical to human LS-DNase. A portion of  
 CC the present sequence was used in Northern blot assays of various murine  
 CC tissues. Expression of murine LS-DNase mRNA was found to be highest in  
 CC the liver and spleen. Human LS-DNase can be used for reducing the  
 CC viscoelasticity or viscous consistency of DNA-containing material, as  
 CC well as treatment of systemic lupus erythematosus or pulmonary disease  
 CC disorders such as cystic fibrosis. It can also be used for adjunctive  
 CC treatment of abscesses or severe closed-space infections, e.g.  
 CC meningitis. It may improve the efficacy of antibiotics used in the  
 CC treatment of such infections (e.g. gentamicin activity is markedly  
 CC reduced by reversible binding to intact DNA). It can also be used in in  
 CC vitro diagnostic assays of a viscous material (e.g. sputum) from a  
 CC patient. The anti-LS-DNase antibodies can be used for detection and  
 CC purification of LS-DNase.  
 CC

SO Sequence 1124 BP; 299 A; 273 C; 273 G; 279 T; 0 other;

Query Match 73.5%; Score 435.2; DB 18; Length 1124;  
 Best Local Similarity 83.4%; Pred. No. 6.9e-128;  
 Matches 494; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 1 GAGAGCTAAACGGAATTCAGAAAAGCATACATACATATGTGATTAGCTTCGC 60  
 DB 404 GAGAGCTGAATGGAATTCAGAGAGACACATATGATGATTAGTTTCGA 463  
 QY 61 CTGGAAGAAACATATTAAGAACAGTATGCTTCTCTATTAAGAAAGCTAGTGTCT 120  
 DB 464 CTGGAAGAAACATATTAAGAACAGTATGCTTCTCTATTAAGAAAGCTAGTGTCT 523  
 QY 121 GTAAAAAAGGACTCTCTATCAAGACTATCAGGCTGAGACGAGATGTGTTCCAGG 180  
 DB 524 GTGAAGACAAATATCACTACATGACTATCAGGATGAGACAGAGCGTGTTCAGG 583  
 QY 181 GAACCTTTGTGTGTGTTCCAGTACCTTACCCGCTGTCAAGAGCTTGTGATGTC 240  
 DB 584 GAACCTTTGTGTGTGTTCCAGTACCTTACCCGCTGTCAAGAGCTTGTGATGTC 643  
 QY 241 CCCTGACACACACCCCTGAGACATCCGTTAGAGATTGATGAGCGGCTGATGTCTAC 300  
 DB 644 CCCTGACACACACCCCTGAGACATCCGTTAAAGAGATGAGTGTGTCGATGTCTAC 703  
 QY 301 ACAGATGTGAAGCTGCTGCTGGAATGACAGAAATTTCTTCAATGCTGATGCT 360  
 DB 704 ACAGATGTGAAGCTGCTGCTGGAATGACAGAAATTTCTTCAATGCTGATGCT 763  
 QY 361 GGTGACAGTACGTCCCAAGAGAGGCTTGAAGACATCCGCTGAGACGAGACCCCAAG 420  
 DB 764 GGTGACAGTACGTCCCAAGAGAGGCTTGAAGACATCCGCTGAGACGAGACCCCAAG 823

QY 421 TTGCTTTGGCTGATCGGGACCAAGAGACACGCTCAAGAAGACAAACTGCGCC 480  
 DB 824 TTGCTTTGGCTGATCGGGACCAAGAGACACGCTCAAGAAGATACAGCTGCGCC 883  
 QY 481 TATGACAGATTCGCTGCTTGAAGACAAATATTTGTCACTCTGTGCTCTCAATCAAC 540  
 DB 884 TATGACAGATTCGCTGCTTGAAGACAAATATTTGTCACTCTGTGCTCTCAATCAAC 943  
 QY 541 CTGCTTTGATTTCCAGAAAGCTTACAGTGTCTGATTCGAAGGCGCTGCG 592  
 DB 944 GCGCTTTGATTTCCAGAAAGCTTACAGTGTCTGATTCGAAGGCGCTGCG 995

# RESULT 13

AAS17902  
 ID AAS17902 standard; cDNA; 350 BP.  
 XX  
 AC AAS17902;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE partial cDNA encoding human fertility-associated antigen (FAA), HC2.  
 XX  
 XX Fertility-associated antigen; FAA; anti-fertility; sperm cell acrosome;  
 KW reproductive fitness; single nucleotide polymorphism; SNP; human;  
 KW HC2; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..349  
 FT /\*tag= a  
 FT /product= "Human FAA"  
 FT /note= "Human fertility-associated antigen, HC2"  
 FT /partial  
 FT /note= "No start or stop codon given"  
 XX  
 PN W0200176529-A2.  
 PD 18-OCT-2001.  
 XX  
 PF 09-APR-2001; 2001WO-US10802.  
 XX  
 PR 07-APR-2000; 2000US-195225P.  
 XX  
 PA (ZHAN/) ZHANG H M.  
 PA (AXRL/) AX R L.  
 XX  
 PI Zhang HM, Ax RL;  
 XX  
 DR WPI; 2002-066307/09.  
 DR P-SDB; AAU11518.  
 XX  
 PT Novel human or bovine fertility-associated antigen useful for  
 PT stabilizing sperm cell acrosome and increasing fertility of a male  
 PS Claim 1; Fig 2A-E; 54pp; English.  
 XX  
 XX The invention describes a novel isolated human fertility-associated  
 CC antigen (FAA) and a bovine FAA polypeptide. Human or bovine FAA is useful  
 CC for increasing the stability of a sperm cell acrosome and for increasing  
 CC the fertility of a human male by administering FAA. Detecting single  
 CC nucleotide polymorphisms (SNPs) of the human or bovine FAA gene is useful  
 CC for assaying fertility by detecting the presence or absence of a FAA SNP  
 CC in the mammal and correlating the presence or absence with the fertility  
 CC of the mammal. The FAA SNPs are chosen from FAA SNPs 1-34 given in the  
 CC specification. FAA SNP is also useful for determining the reproductive  
 CC fitness of a human or bovine, by determining the nucleotide sequence of  
 CC the FAA SNP, quantifying the fertility of more than one mammal containing  
 CC the FAA SNP and correlating the frequency of the FAA SNP to the  
 CC reproductive fitness of the mammals. FAA improves the integrity of sperm  
 CC membranes and increases the capacitation of sperm derived from either  
 CC fertile or infertile humans. Early genetic identification of infertility

CC improves the prognosis for subsequent attempts at fertilisation and  
CC facilitates early intervention to determine whether the individual may  
CC benefit from fertility treatments to avoid costly and/or emotional  
CC problems with attempted inseminations. This sequence encodes the human  
CC fertility-associated antigen (FAA), HC3, one of 4 partial prostate cDNA  
CC clones of the novel FAA gene described in the method of the invention.  
XX

SQ Sequence 350 BP; 88 A; 91 C; 93 G; 78 T; 0 other;

Query Match 58.9%; Score 348.4; DB 24; Length 350;  
Best Local Similarity 99.7%; Pred. No. 1.6e-100;  
Matches 349; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 177 CAGGAACCCCTTTGGTCTGGTCCAGTACCCCTACACCGCTGTCAAGGACTTCGTGAT 236  
Db 1 CAGGAACCCCTTTGGTCTGGTCCAGTACCCCTACACCGCTGTCAAGGACTTCGTGAT 60  
QY 237 TGTCCCTCGACACACCCCTGTGAGATCGTTAGAGAGATTGATGAGCTGGCTGATGT 296  
Db 61 TGTCCCTCGACACACCCCTGTGAGATCGTTAGAGAGATTGATGAGCTGGCTGATGT 120  
QY 297 CTACACAGATGTGAACCGTCTCGAATGTCAGAGAAATTTTCAATTTTCATGGTGACTTCAA 356  
Db 121 CTACACAGATGTGAACCGTCTCGAATGTCAGAGAAATTTTCAATTTTCATGGTGACTTCAA 180  
QY 357 TGCTGGCTGCAGCTACGTCCCAAGAGCGCTGGAAGACATCCGCTGAGGACGGACCC 416  
Db 181 TGCTGGCTGCAGCTACGTCCCAAGAGCGCTGGAAGACATCCGCTGAGGACGGACCC 240  
QY 417 CAAGTTCTGTTGGCTGATCGGGGACCAAGAGGACACACCGTCAAGAGACCAAACTG 476  
Db 241 CAAGTTCTGTTGGCTGATCGGGGACCAAGAGGACACACCGTCAAGAGACCAAACTG 300  
QY 477 CGCCTATGACAGGATCGTCTTAGAGGACAAATATTGTCAACTCTGGT 526  
Db 301 CGCCTATGACAGGATCGTCTTAGAGGACAAATATTGTCAACTCTGGT 350

## RESULT 14

AAS17903  
ID AAS17903 standard; cDNA; 350 BP.

AC AAS17903;

XX 26-MAR-2002 (first entry)

DE partial cDNA encoding human fertility-associated antigen (FAA), HC3.

XX Fertility-associated antigen; FAA; antiinfertility; sperm cell acrosome;  
KW reproductive fitness; single nucleotide polymorphism; SNP; human;  
XX HC3; ss.

OS Homo sapiens.

FH Key Location/Qualifiers

FT CDS 2..349

FT /\*tag= a

FT /product= "Human FAA"

FT /note= "Human fertility associated antigen, HC3"

FT /partial

FT /note= "No start or stop codon given"

FT /transl\_except= (pos:251..253, aa: Arg)

XX WO200176529-A2.

XX 18-OCT-2001.

XX 09-APR-2001; 2001WO-US10802.

XX 07-APR-2000; 2000US-195225P.

XX (ZHAN/) ZHANG H. M.

PA (AXRL/) AX R. L.

XX

PI Zhang HM, Ax RL;

DR WPI; 2002-066307/09.

DR P-PSDB; AAU11519.

XX Novel human or bovine fertility-associated antigen useful for  
PT stabilising sperm cell acrosome and increasing fertility of a male -

XX Claim 1; Fig 2A-E; 54pp; English.

XX The invention describes a novel isolated human fertility-associated  
CC antigen (FAA) and a bovine FAA polypeptide. Human or bovine FAA is useful  
CC for increasing the stability of a sperm cell acrosome and for increasing  
CC the fertility of a human male by administering FAA. Detecting single  
CC nucleotide polymorphisms (SNPs) of the human or bovine FAA gene is useful  
CC for assaying fertility by detecting the presence or absence of a FAA SNP  
CC in the mammal and correlating the presence or absence with the fertility  
CC of the mammal. The FAA SNPs are chosen from FAA SNPs 1-34 given in the  
CC specification. FAA SNP is also useful for determining the reproductive  
CC fitness of a human or bovine, by determining the nucleotide sequence of  
CC the FAA SNP, quantifying the fertility of more than one mammal containing  
CC the FAA SNP and correlating the frequency of the FAA SNP to the  
CC reproductive fitness of the mammals. FAA improves the integrity of sperm  
CC membranes and increases the capacitation of sperm derived from either  
CC fertile or infertile humans. Early genetic identification of infertility  
CC improves the prognosis for subsequent attempts at fertilisation and  
CC facilitates early intervention to determine whether the individual may  
CC benefit from fertility treatments to avoid costly and/or emotional  
CC problems with attempted inseminations. This sequence encodes the human  
CC fertility-associated antigen (FAA), HC3, one of 4 partial prostate cDNA  
CC clones of the novel FAA gene described in the method of the invention.

XX SQ Sequence 350 BP; 89 A; 93 C; 94 G; 74 T; 0 other;

Query Match 53.2%; Score 314.8; DB 24; Length 350;  
Best Local Similarity 93.7%; Pred. No. 7.9e-90;  
Matches 328; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 177 CAGGGAACCCCTTTGTGCTGTGGTCCAGTACCCCTACACCGCTGTCAAGGACTTCGTGAT 236

Db 1 CAGGGAACCCCTTTGTGCTGTGGTCCAGTACCCCTACACCGCTGTCAAGGACTTCGTGAT 60

QY 237 TGTCCCTCGACACACCCCTGTGAGATCGTTAGAGAGATTGATGAGCTGGCTGATGT 296

Db 61 TATCCCTCGACACACCCCTGTGAGATCGTTAGAGAGATTGATGAGCTGGCTGATGT 120

QY 297 CTACACAGATGTGAACCGTCTGGAATGTCAGAGAAATTTTCAATTTTCATGGTGACTTCAA 356

Db 121 CTACACAGATGTGAACCGTCTGGAATGTCAGAGAAATTTTCAATTTTCATGGTGACTTCAA 180

QY 357 TGCTGGCTGCAGCTACGTCCCAAGAGCGCTGGAAGGACATCCGCTGAGGACGGACCC 416

Db 181 TGCTGGCTGCAGCTACGTCCCAAGAGCGCTGGAAGGACATCCGCTGAGGACGGACCC 240

QY 417 CAAGTTCTGTTGGCTGATCGGGGACCAAGAGGACACACCGTCAAGAGACCAAACTG 476

Db 241 CAAGTTCTGTTGGCTGATCGGGGACCAAGAGGACACACCGTCAAGAGACCAAACTG 300

QY 477 CGCCTATGACAGGATCGTCTTAGAGGACAAATATTGTCAACTCTGGT 526

Db 301 CGCCTATGACAGGATCGTCTTAGAGGACAAATATTGTCAACTCTGGT 350

## RESULT 15

AAS17904

ID AAS17904 standard; cDNA; 349 BP.

XX AAS17904;

XX 26-MAR-2002 (first entry)

DE Partial cDNA encoding human fertility-associated antigen (FAA), HC4.

XX	Fertility-associated antigen; FAA; antiinfertility; sperm cell acrosome;	
KW	reproductive fitness; single nucleotide polymorphism; SNP; human;	
XX	HC4; ss.	
OS	Homo sapiens.	
XX		
FT	Key	Location/Qualifiers
FT	CD8	2..348
FT		/tag= a
FT		/product= "Human FAA"
FT		/note= "Human fertility-associated antigen, HC4"
FT		/partial
FT		/note= "No start or stop codon given"
FT		/transl except= (pos: 11..12, aa:Phe)
FT		/note= "This codon has an apparent 1 nucleotide deletion which alters the reading frame"
XX		
PN	WO200176529-A2.	
XX		
PD	18-OCT-2001.	
XX		
PF	09-APR-2001; 2001WO-US10802.	
XX		
PR	07-APR-2000; 2000US-195225P.	
XX		
PA	(ZHAN/) ZHANG H. M.	
XX	(AXRL/) AX R L.	
P1		
XX	Zhang HM, Ax RL;	
DR	WPI; 2002-066307/09.	
XX	P-PSDB; AAU11520.	
PT	Novel human or bovine fertility-associated antigen useful for	
XX	stabilising sperm cell acrosome and increasing fertility of a male	
XX	Claim 1; Fig 2A-E; 54pp; English.	
XX		
CC	The invention describes a novel isolated human fertility-associated	
CC	antigen (FAA) and a bovine FAA polypeptide. Human or bovine FAA is useful	
CC	for increasing the stability of a sperm cell acrosome and for increasing	
CC	the fertility of a human male by administering FAA. Detecting single	
CC	nucleotide polymorphisms (SNPs) of the human or bovine FAA gene is useful	
CC	for assaying fertility by detecting the presence or absence of a FAA SNP	
CC	in the mammal and correlating the presence or absence with the fertility	
CC	of the mammal. The FAA SNPs are chosen from FAA SNPs 1-34 given in the	
CC	specification. FAA SNP is also useful for determining the reproductive	
CC	fitness of a human or bovine. By determining the nucleotide sequence of	
CC	the FAA SNP, quantifying the fertility of more than one mammal containing	
CC	the FAA SNP and correlating the frequency of the FAA SNP to the	
CC	reproductive fitness of the mammals. FAA improves the integrity of sperm	
CC	membranes and increases the capacitation of sperm derived from either	
CC	fertile or infertile humans. Early genetic identification of infertility	
CC	improves the prognosis for subsequent attempts at fertilisation and	
CC	beneficiates early intervention to determine whether the individual may	
CC	benefit from fertility treatments to avoid costly and/or emotional	
CC	problems with attempted inseminations. This sequence encodes the human	
CC	fertility-associated antigen (FAA). HC4, one of 4 partial prostate cDNA	
CC	clones of the novel FAA gene described in the method of the invention.	
XX		
Q0	Sequence 349 BP; 89 A; 91 C; 93 G; 76 T; 0 other;	

Query Match	50.8%	Score 300.6;	DB 24;	Length 349;
Best Local Similarity	91.6%	Pred. No. 2.6e-85;		
Matches 318; Conservative	0;	Mismatches 29;	Indels 0;	Gaps 0

QY 180 GGAACCCCTTGGGCTCTGTTCCAGTCAACCTAACCGCGTCCAAAGACTTCGGATTGT 23  
Db 3 GGGAGCCCTTGTGGTCTGTCCAAATCTCCCCAACCTGCTCAAGACTTCGGATTAT 62  
QY 240 CCCCCTGCACACCAACCCCTGAGACATCCGTTAGAGAGATTATGACCTGCGTGATGTCTA 29

Db	63	CCCCCTGACACCCACCCAGACATCCGTTAAGGAGATGATGAGTTGGTTAGGCTA	1222
Qy	300	CACAGATGTGAAAGAGTGGCTGGGAATGCAGAGAAATTTATTTTCATGGGTGACTTCAATGC	3588
Db	123	CACGAGCGTGAACACCGCTGGAGAGGGGAGAAATTTATTTTCATGGGTGACTTCAATGC	1820
Qy	360	TGGCTGCAGCTACGTCGCCAAGAAAGCGCTGGAAGACATCCGCTTGAGAGACGAGCCCA	4191
Db	183	CGGCTGCAGCTACGTCGCCAAGAAAGCGCTGGAAGAAATCCGCTTGAGAGCTGACCCAG	2424
Qy	420	GTTTGGTTGGCTGATCGGGGACCAAGAGACACCAAGCTCAAGAGAGCAATACTGGC	4797
Db	243	GTTTGGTTGGCTGATCGGGGACCAAGAGAGACACCAAGCTCAAGAGAGCAATACTGGC	3020
Qy	480	CTATGACAGGATCGTGGCTTGAAGAGCAAAATATTGTCAATCTGGGTG	526
Db	303	CTATGACAGGATCGTGGCTTGAAGAGCAAAATATTGTCAATCTGGGTG	349

Search completed: December 9, 2002, 09:17:55  
Job time : 179.616 secs

**inis Page Blank (uspto)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 09:20:03 ; Search time 1320.4 Seconds  
(without alignments)  
7261.235 Million cell updates/sec

Title: US-09-905-114-1  
Perfect score: 592  
Sequence: 1 ggaagctaaacggaatc.....gctgaatcgaagccctcg 592

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gse:\*  
18: em\_gse\_hum:\*  
19: em\_gse\_inv:\*  
20: em\_gse\_pin:\*  
21: em\_gse\_vrt:\*  
22: em\_gse\_fun:\*  
23: em\_gse\_mam:\*  
24: em\_gse\_mus:\*  
25: em\_gse\_other:\*  
26: em\_gse\_pro:\*  
27: em\_gse\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	478.4	80.8	890	9	AL546857 AL546857
2	476.8	80.5	896	9	AL546432 AL546432
3	471.8	79.7	833	9	AL572037 AL572037
4	469	79.2	931	9	AL571894 AL571894
5	438.8	74.1	506	10	AV607154 AV607154
6	432	73.0	666	10	BE048177 BE048177

C 7	421	71.1	695	13	B1522352	603081420
C 8	411.6	69.5	642	10	AM543356	CO168H07
C 9	406	68.6	825	12	BG867772	BG867772
10	396.4	67.0	825	12	BG530136	BG530136
C 11	391.2	66.1	820	10	AM543802	AM543802
C 12	389.4	65.8	628	13	B1460950	B1460950
C 13	386	65.2	610	10	AM540389	AM540389
C 14	383.8	64.8	668	10	AM915564	AM915564
C 15	383.6	64.8	784	9	A1927844	A1927844
C 16	376	63.5	832	13	B1838188	B1838188
C 17	363.4	61.4	634	9	A1660448	A1660448
C 18	363.4	61.4	981	14	B0710632	B0710632
C 19	361	61.0	685	13	B1833952	B1833952
C 20	353.2	59.7	566	10	AM538998	AM538998
C 21	351	59.3	759	12	BG534290	BG534290
C 22	330.2	55.8	527	10	AM539589	AM539589
C 23	328.2	55.4	613	10	AM779563	AM779563
C 24	328	55.4	578	10	AM658627	AM658627
C 25	311	52.5	643	10	AM001760	AM001760
C 26	303.8	51.3	714	12	BG566893	BG566893
C 27	302.2	51.0	855	12	BF237166	BF237166
C 28	296.8	50.1	542	12	BF434855	BF434855
C 29	291.2	49.2	901	12	BF303508	BF303508
C 30	290	49.0	554	10	AM243180	AM243180
C 31	288.8	48.8	555	9	A134430	A134430
C 32	288.2	48.7	487	9	A1194886	A1194886
C 33	279.8	47.3	341	10	AM803005	AM803005
C 34	279.6	47.2	497	9	AA564572	AA564572
C 35	277.2	46.8	586	13	BM489658	BM489658
C 36	267.6	45.2	522	14	BQ186387	BQ186387
C 37	265.6	44.9	595	12	BG077563	BG077563
C 38	262.6	44.4	453	14	H78170	H78170
C 39	261.2	44.1	433	12	BF911384	BF911384
C 40	252.4	42.6	511	9	A1271665	A1271665
C 41	248.4	42.0	523	9	A1248904	A1248904
C 42	240.4	40.6	536	9	AA707043	AA707043
C 43	239.6	40.5	486	9	AA989198	AA989198
C 44	237.6	40.1	498	12	BG064211	BG064211
C 45	236.2	39.9	427	10	AM827162	AM827162

#### ALIGNMENTS

RESULT 1  
LOCUS AL546857 890 bp mRNA linear EST 16-FEB-2001  
DEFINITION AL546857 LTI\_NFL006\_PL2 Homo sapiens cDNA clone CSOD1026YK08 5  
prime, mRNA sequence.  
ACCESSION AL546857  
VERSION AL546857.1 GI:12880381  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 890)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
FEATURES  
source  
1..890  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CSOD1026YK08"  
/clone\_lib="LTI\_NFL006\_PL2"  
/tissue\_type="Placenta"  
/note="Vector: PCWVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end



enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/

BASE COUNT 250 a 215 c 220 g 204 t 1 others

Query Match 80.8%; Score 478.4; DB 9; Length 890;  
Best Local Similarity 88.0%; Pred. No. 1.1e-128;  
Matches 521; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 1 GAGAGCTTAAACGGAATTCAGAGAAAGGCAATACATACACTATGTGATTAGCTCTCGC 60  
Db 216 GAGAGCTTAAACGGAATTCAGAGAAAGGCAATACATACACTATGTGATTAGCTCTCG 275  
QY 61 CTGGAAGAAACACATATAAAGACAGTATGCTTTCTCTATAAGAAAGCTAGTGTCT 120  
Db 276 CTGGAAGAAACACATATAAAGACAGTATGCTTTCTCTATAAGAAAGCTAGTGTCT 335  
QY 121 GTAAACAAAGCTACCTTACCAGACTATCAGCTGGAGAGCAGATGTGTTTCCAGG 180  
Db 336 GTGAAGAGGAGTTATCACTACCATGACTATCAGATGGAGAGCAGATGTGTTTCCAGG 395  
QY 181 GAACCTTTGTGCTGCTTCCAGTCACTACACCGCTGTCAAGGACTTCGTTGATGTCT 240  
Db 396 GAGCCCTTTGTGCTGCTTCCAGTCACTACACCGCTGTCAAGGACTTCGTTGATGTCT 455  
QY 241 CCCCTGACACACCCCTGAGACATCCGTTAGAGAGATGTGAGCTGGCTGATGTCTAC 300  
Db 456 CCCCTGACACACCCCTGAGACATCCGTTAGAGAGATGTGAGCTGGCTGATGTCTAC 515  
QY 301 ACAGATGTGAACGTCGCTGGAATCAGAGAAATTTCAATTTTCATGGTGACTTCAATGCT 360  
Db 516 ACGGACGTGAACACACCGCTGGAAGCGGAGAAATTTCAATTTTCATGGTGACTTCAATGCT 575  
QY 361 GGCTGAGCTAGCTGCTCCCAAGAGCCCTGGAAGGACATCCGCTGAGAGACGCCCAAG 420  
Db 576 GGCTGAGCTAGCTGCTCCCAAGAGCCCTGGAAGGACATCCGCTGAGAGACGCCCAAG 635  
QY 421 TTCGTTTGGCTGATCGGGACCAAGAGACACCGCTCAAGAGAGACAACTGGCC 480  
Db 636 TTTGTTTGGCTGATCGGGACCAAGAGACACCGCTCAAGAGAGACAACTGGCC 695  
QY 481 TATGACAGGATCGCTTATAGAGCAAAATATTTGTCAGTCTGCTGCTCAATCAAAAC 540  
Db 696 TATGACAGGATCGCTTATAGAGCAAAATATTTGTCAGTCTGCTGCTCAATCAAAAC 755  
QY 541 CTCGCTTTTGTATTCAGAAAGCTTACAGGTTGTCTGTAATCAAGGCGCTCG 592  
Db 756 AGTGTGTTTGTATTCAGAAAGCTTACAGGTTGTCTGTAATCAAGGCGCTCG 807

RESULT 2  
AL546432 896 bp mRNA linear EST 16-FEB-2001  
LOCUS AL546432 LTI\_NFL006\_PL2 Homo sapiens cDNA clone CS0D1030YD20 5  
DEFINITION prime, mRNA sequence.

ACCESSION AL546432  
VERSION AL546432.1 GI:12879540  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1. (bases 1 to 896)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES  
source

1..896  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0D1030YD20"  
/clone\_lib="LTI\_NFL006\_PL2"  
/tissue\_type="placenta"  
/note="Vector: pCMVSPORT 6; Site: 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/

BASE COUNT 249 a 217 c 226 g 203 t 1 others

Query Match 80.5%; Score 476.8; DB 9; Length 896;  
Best Local Similarity 87.8%; Pred. No. 3.1e-128;  
Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 1 GAGAGCTTAAACGGAATTCAGAGAAAGGCAATACATACACTATGTGATTAGCTCTCGC 60  
Db 266 GAGAGCTTAAACGGAATTCAGAGAAAGGCAATACATACACTATGTGATTAGCTCTCG 325  
QY 61 CTGGAAGAAACACATATAAAGACAGTATGCTTTCTCTATAAGAAAGCTAGTGTCT 120  
Db 326 CTGGAAGAAACACATATAAAGACAGTATGCTTTCTCTATAAGAAAGCTAGTGTCT 385  
QY 121 GTAAACAAAGCTACCTTCAACGACTATCAGCTGGAGACGACAGATGTGTTTCCAGG 180  
Db 386 GTGAAGAGGAGTTATCACTACCATGACTATCAGATGGAGACGACAGATGTGTTTCCAGG 445  
QY 181 GAACCCCTTTGTGCTGCTGCTTCCAGTCACTACACCGCTGTCAAGGACTTTCGTTGTTGTC 240  
Db 446 GAGCCCTTTGTGCTGCTGCTTCCAGTCACTACACCGCTGTCAAGGACTTTCGTTGTTGTC 505  
QY 241 CCCCTGACACACCCCTGAGACATCCGTTAGAGAGATGTGAGCTGGCTGATGTCTAC 300  
Db 506 CCCCTGACACACCCCTGAGACATCCGTTAGAGAGATGTGAGCTGGCTGATGTCTAC 565  
QY 301 ACAGATGTGAACGTCGCTGGAATGAGAGAAATTTCAATTTTCATGGTGACTTCAATGCT 360  
Db 566 ACGGACGTGAACACACCGCTGGAAGCGGAGAAATTTCAATTTTCATGGTGACTTCAATGCT 625  
QY 361 GGCTGAGCTACGTCCTCCCAAGAGGCTTGAAGAGACATCCGCTGAGAGCGGACCCCAAG 420  
Db 626 GGCTGAGCTACGTCCTCCCAAGAGGCTTGAAGAGACATCCGCTGAGAGCTGACCCCAAG 685  
QY 421 TTCGTTTGGCTGATCGGGACCAAGAGGACACCCGCTCAAGAGAGACAAACTGCGCC 480  
Db 686 TTTGTTTGGCTGATCGGGACCAAGAGGACACCCGCTCAAGAGAGACAAACTGCTGCA 745  
QY 481 TATGACAGGATCGCTTATAGAGCAAAATATTTGTCAGTCTGCTGCTCAATCAAAAC 540  
Db 746 TATGACAGGATCGCTTATAGAGCAAAATATTTGTCAGTCTGCTGCTCAATCAAAAC 805  
QY 541 CTCGCTTTTGTATTCAGAAAGCTTACAGGTTGTCTGTAATCAAGGCGCTCG 592  
Db 806 AGTGTGTTTGTATTCAGAAAGCTTACAGGTTGTCTGTAATCAAGGCGCTCG 857

RESULT 3  
AL572037/c  
LOCUS AL572037 LTI\_NFL006\_PL2 Homo sapiens cDNA clone CS0D1026YK08 3  
DEFINITION prime, mRNA sequence.

AL572037 833 bp mRNA linear EST 16-FEB-2001  
LOCUS AL572037 LTI\_NFL006\_PL2 Homo sapiens cDNA clone CS0D1026YK08 3  
DEFINITION prime, mRNA sequence.

Accession	Version	Source	Organism	Reference	Authors	Title	Journal	Comment
AL572037	AL572037.1	GI:12929919	human.	Human sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	1 (bases 1 to 833)	Li, W.B., Gruber, C., Jeessee, J., and Polayes, D.	Full-length cDNA libraries and normalization unpublished (2001)
BP 191 91006	ENRY cedex - France	Email: seque@genoscope.cns.fr, web : www.genoscope.cns.fr.	Genoscope - Centre National de Sequencage					
Location/Qualifiers								
1..833								
/organism="Homo sapiens"								
/db_xref="taxon:9606"								
/clone="CS0D1026YK08"								
/clone_lib="LTI_NFL006_PL2"								
/issue_type="Placenta"								
/note="Vector: pcwvSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo (GT) primer. Five prime and enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pcwvSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Peng liang life technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email: fliang@lifetech.com URL : http://fulllength.invitrogen.com"								
BASE COUNT	193 a	193 c	191 g	249 t	7 others			
ORIGIN								
Query Match	79.7%	Score 471.8	DB 9	Length 833				
Best Local Similarity	87.0%	Pred. No. 8.7e-127						
Matches 515	Conservative 2	Mismatches 75	Indels 0	Gaps 0				
1	GAGAACTTAAACGAAATTCAGAAAGGATCAATCACTATGTGATTTAGCTTGGC							
734	GAGAACTTAAACGAAATTCAGAAAGGATCAATCACTATGTGATTTAGCTTGGC							
61	CTTGGAAGAAACATATTAAGACAGTATGCTTTCTTATTAAGAAAGCTAGTGTCT							
674	CTTGGAAGAAACATATTAAGACAGTATGCTTTCTTATTAAGAAAGCTAGTGTCT							
121	GTAAGAAACAGTACTTACCTACCGACATATCAGCTGAGAGAGCGAGATGTCTTCCAGG							
614	GTAAGAGAGATTTACCTACCTACCTATCAGATGAGAGAGAGATGTCTTCCAGG							
181	GAACCTTTGTGTGTGTGTTCAGTACACCGCTGATCAAGAACTTGTGATTTGTC							
554	GAGCCCTTTGTGTGTGTTCAGTACACCGCTGATCAAGAACTTGTGATTTGTC							
241	CCCTGACACCAACCCCTGAGACATCCGTTAGAGAGATTTGATGAGCTGATGATCTAC							
494	CCCTGACACCAACCCCTGAGACATCCGTTAGAGAGATTTGATGAGCTGATGATCTAC							
301	ACAGATGTGAAACGTCGCTGGAATGCAAGAAATTTTCAATTTTATGGGTGACTTCAATGCT							
434	ACGAGCGTGAACACCGCTGGAAGCGGGAATTTTCAATTTTATGGGTGACTTCAATGCT							
361	GAGTCACACTAGTCCCAAGAAAGCTTGAAGACATCCGCTGAGAGCGGAGCCCAAG							
374	GAGTCACACTAGTCCCAAGAAAGCTTGAAGACATCCGCTGAGAGCGGAGCCCAAG							
421	TTCTGTTGGCTGATCGGGGACCAAGAGGACACACGCTTAAGAGAGCAAACTGCGCC							
314	TTTGTGTTGGCTGATCGGGGACCAAGAGGACACACGCTTAAGAGAGCAAACTGCGCA							
481	TATGACAGGATCGTCTTAGAGACAAATATTTGTCACCTCTGTGTGTCTCAATCAAC							

Db	254	TATAGACAGATTGGCTTAGAGGACAGAAATCGCAGCTTCTGTTGTTCCAGATCAAC	195
Qy	541	CTCGCTTTGATTTTCAGAAAGCTTACAGGTTGTCTGATGAAGCCCTGG	592
Db	194	AGTGTTTTGACTTCAGAAAGCTTACAAAGCTGACTGAAGAGAGGCCCTGG	143
RESULT 4			
LOCUS	AL571894/c	931 bp	mRNA
DEFINITION	AL571894 LTI NFL006 PL2 Homo sapiens cDNA clone CS0D1030YD20 3		linear EST 16-FEB-2001
ACCESSION	AL571894		
VERSION	AL571894.1	GI:12929641	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryotic: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Li, W.-B., Gruber, C., Jesse, J. and Polayes, D.		
TITLE	Full-length cDNA libraries and normalization		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr. Location/Qualifiers		
FEATURES	1..931		
source	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="CS0D1030YD20" /clone_1b="LTI NFL006 PL2" /issue_type="placenta" /note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@life tech.com URL : http://fulllength.invitrogen.com"		
BASE COUNT	209 a	219 c	225 g
ORIGIN			271 t
			7 others
Query Match	79.2%	Score 469:	DB 9: Length 931:
Best Local Similarity	86.8%	Pred. No. 6e-126:	
Matches 514: Conservative	1: Mismatches 77:	Indels 0:	Gaps 0:
Qy	1	GAGAAAGCTAAACGGAATTCACAAGAAAGGCTTAACATACAACTATGATAGCTTCGC	60
Db	726	GAGAAAGCTGAACAGAAATTCAGAGAGAGGCTTAACGTAACACTATGATAGCTTCGG	667
Qy	61	CTTGAAGAAAACACATATAAAGAACAGATGCTCTTCTCTATAAAGAAAGCTAGTGCT	120
Db	666	CTTGAAGAAAACACATATAAAGAACATATATGCTTCTCTACAAGAAAGCTGATGCT	607
Qy	121	GTAACAAAGCTACCTCTACACGACTATATAGGCTGGAGACGAGATGTTTTCCAGG	180
Db	606	GTAAGAGAGATTTATACATACATGACTATAGATGAGAGACGAGATGTTTTCCAGG	547
Qy	181	GAACCTTTGTGTCTGCTGATTCAGTACACCCCTACACCGCTGCAAGAACTTGATTTGC	240
Db	546	GAGCCCTTTGTGTCTGCTGATTCATCTCCCAACACTGCTGCAAGAACTTGATTTATC	487
Qy	241	CCCTGACACACACCTCTGAGACATCGGTATAGAGAGATGATGAGCTGATGCTTAC	300
Db	486	CCCTGACACACACCTCTGAGACATCGGTATAGAGATGATGATGATGATGCTTAC	427
Qy	301	ACAAGTGAAGCTCGCTGGAATGCAAGAAATTTCAATTTTCAATGAGTTCAATGCT	360

```
Db 426 ACGACGTGAACACCCTGGAAGCGGAGAAATTTCAATTTTCATGGGTGACTCAATGCC 367
QY 361 GGCTGCAGCTACGTCCTCCCAAGAGCCCTGGAAAGGACATCCCGCTTGAGGACGGACCCCAAG 420
Db 366 GGCTGCAGCTACGTCCTCCCAAGAGCCCTGGAAAGGACATCCCGTTGAGAACTGACCCCCAGG 307
QY 421 TTCGTTTGGCTGATCGGGGACCAAGAGACACCAACGGTCAAGAAAGACACAACTGGGCC 480
Db 306 TTTGTTTGGCTGATCGGGGACCAAGAGACACCAACGGTCAAGAAAGACACCAACTGTGCA 247
QY 481 TATGACAGGATCGTCTTATAGAGGACAAATATTCTCAACTCTGCTGCTCTCAATCAAAAC 540
Db 246 TATAACAGGATGTGCTTATAGAGGACAAATATTCTCAACTCTGCTGCTCTCAATCAAAAC 187
QY 541 CTCGCTTTGATTTCCAGAAAGCTTACAGGTGTCTGAATCGAAGGCGCCTGG 592
Db 186 AGTGTTTTGACTCCAGAAAGCTTACAGCTGACTGAAGAGGAGGCCCTGG 135

RESULT 5
AV607154/c
LOCUS AV607154 Bos taurus kidney fetus Bos taurus cDNA clone EST 28-NOV-2001
DEFINITION 3', mRNA sequence.
ACCESSION AV607154
VERSION AV607154.1 GI:9737527
KEYWORDS EST.
SOURCE COW.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 506)
AUTHORS Takasuga,A., Hirotsune,S., Itoh,R., Jitohzono,A., Suzuki,H., Aso,H.
and Sugimoto,Y.
TITLE Establishment of a high throughput EST sequencing system using
poly(A) tail-removed cDNA libraries and determination of 36,000
bovine ESTs
JOURNAL Nucleic Acids Res. 29 (22), E108 (2001)
MEDLINE 21570554
COMMENT Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shrakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazusugi@cooca.ocn.ne.jp
Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.
FEATURES
source
1..506
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="E1KI039B05"
/clone_lib="Bos taurus kidney fetus"
/tissue_type="kidney"
/dev_stage="fetus"
/lab_host="DH10B"
/notes="Vector: pZL1; Site_1: SalI; Site_2: NotI; Poly A
was deleted from a NotI site"
BASE COUNT 124 a 126 c 129 g 126 t 1 others
ORIGIN
Query Match 74.1%; Score 438.8; DB 10; Length 506;
Best Local Similarity 99.3%; Pred. No. 3.1e-117;
Matches 440; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 150 TCAGGTGAGACCGCAGATGTGTTTTCAGGGACCCCTTGTGGTCCAGTCACCTCACC 209
Db 506 TCAGGTGAGACCGCAGATGTGTTTTCAGGGACCCCTTGTGGTCCAGTCACCTCACC 447
QY 210 CTACACCGCTGTCAAGGACTTCGTGATTTGTCCTCCCTGCACACCCCTTGAGACATCCGT 269
```

```
Db 446 CTACACCGCTGTCAAGGACTTNGTGTGTCCTCCCTGCACACCCCTTGAGACATCCGT 387
QY 270 TAGAGAGATTGATGAGCTGGCTGTGATGTCTACACAGATGTGAACCTCGCTGGAATGCAGA 329
Db 386 TAGAGAGATTGATGAGCTGGCTGTGATGTCTACACAGATGTGAACCTCGCTGGAATGCAGA 327
QY 330 GAATTTTCATTTTCATGGGTGACTTCAATGCTGGCTGACGTACGTCCTCCCAAGAGCCCTG 389
Db 326 GAATTTTCATTTTCATGGGTGACTTCAATGCTGGCTGACGTACGTCCTCCCAAGAGCCCTG 267
QY 390 GAAGGACATCCCGCTCGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGA 449
Db 266 GAAGGACATCCCGCTCGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGA 207
QY 450 CACCACGGTCAAGAGAGAGACACAACTGCCCTATGACAGGATCGTGTCTTAGAGGACAAAA 509
Db 206 CACCACGGTCAAGAGAGAGACACAACTGCCCTATGACAGGATCGTGTCTTAGAGGACAAAA 147
QY 510 TATTGTCAACTCTGGTGGTCTCAATCAAACTCGTCTTTGATTTCCAGAAAGCTTTACAG 569
Db 146 TATTGTCAACTCTGGTGGTCTCAATCAAACTCGTCTTTGATTTCCAGAAAGCTTTACAG 87
QY 570 GTTGTCTGAATCGAAGGCCCTGG 592
Db 86 GTTGTCTGAATCGAAGGCCCTGG 64

RESULT 6
BE048177
LOCUS BE048177
DEFINITION 666 bp mRNA linear EST 20-OCT-2000
similar to SW:DH22 HUMAN Q13609 DNASE I HOMOLOGOUS PROTEIN DHP2
PRECURSOR ;, mRNA sequence.
ACCESSION BE048177
VERSION BE048177.1 GI:8365155
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 666)
AUTHORS NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BRGAP), Tumor Gene Index
JOURNAL Unpublished (1998)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Insert Length: 1048 Std Error: 0.00
Seq primer: -40RP from Gibco
High quality sequence stop: 431.
Location/Qualifiers
1..666
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2291681"
/clone_lib="NCI_CGAP_Brn52"
/tissue_type="tumor, 5 pooled (see description)"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; This library represents the normalized
version of NCI CGAP Brn35. Cloned unidirectionally.
Primer: Oligo dt. Average insert size 1.19 kb. Tumor
types include: meningioma, oligodendroglioma, astrocytoma
(grade II), medulloblastoma, astrocytoma (grade IV)."
```

BASE COUNT 198 a 149 c 170 g 147 t 2 others  
ORIGIN

Query Match 73.0%; Score 432; DB 10; Length 666;  
Best Local Similarity 87.5%; Pred. No. 3,4e-115;  
Matches 471; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 1 GAGAGCTAAACGAAATTCAGAAAGGACATACATACATATGATTAAGCTTCGC 60  
DB 121 GAGAGCTAAACGAAATTCAGAAAGGACATACATACATATGATTAAGCTTCGC 180  
QY 61 CTGGAAGAAACACATATTAAGAAACAGTATGCTTTCTATATAAGAAAGCTAGTCT 120  
DB 181 CTGGAAGAAACACATATTAAGAAACATATATGCTTTCTATACAGAAAGCTAGTCT 240  
QY 121 GTAAACAAAGCTACCTCTACACAGACTATATAGGCTGAGACGCAATGTGTTTCCAGG 180  
DB 241 GTAAAGAGAGTATACATACATACATACATACATACATACATACATACATACAT 300  
QY 181 GAACCTTTTGTGTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
DB 301 GAGCCCTTTGT 360  
QY 241 CCCCTGACACACACCCCTGACACATCCGTTAGAGATTGATGATGATGATGATGAT 300  
DB 361 CCCCTGACACACACCCCTGACACATCCGTTAGAGATTGATGATGATGATGATGAT 420  
QY 301 ACAGATGTGAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
DB 421 ACAGATGTGAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
QY 361 GGCTGACGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
DB 481 GGCTGACGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
QY 421 TTGCTTTGCTGATCGGAGACCAAGAGACACCAAGTTCAGAAAGACCAACATGTGCA 480  
DB 541 TTGCTTTGCTGATCGGAGACCAAGAGACACCAAGTTCAGAAAGACCAACATGTGCA 600  
QY 481 TATGACAGGATGCTGCTTGAAGGACCAAGAAATGCAAGTTCGTTCCCAAGTCAA 538  
DB 601 TATGACAGGATGCTGCTTGAAGGACCAAGAAATGCAAGTTCGTTCCCAAGTCAA 658

RESULT 7  
B1522352/c 695 bp mRNA linear EST 29-AUG-2001  
LOCUS 603081420T1 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:520532 3',  
DEFINITION mRNA sequence.  
ACCESSION B1522352  
VERSION B1522352.1 GI:15347144  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 695)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM11554 row: h column: 05  
High quality sequence start: 18  
High quality sequence stop: 695.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

FEATURES  
source Location/Qualifiers

1. 695  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:520532"  
/clone\_11b="NIH\_MGC\_120"  
/lab\_host="DH10B"  
/note="Organ: pooled pancreas and spleen; Vector:  
pCMV-Sport6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA  
source anonymous pool of spleen and pancreas from 28 yo  
male. Library is oligo-dT primed and directionally cloned  
(EcoRV site is destroyed upon cloning). Average insert  
size 1.5 kb, insert size range 1-2.5 kb. Library is  
normalized and enriched for full-length clones and was  
constructed by C. Gruber (Invitrogen) Research Genetics  
tracking code 025. Note: this is a NIH\_MGC Library."

BASE COUNT 159 a 171 c 163 g 202 t  
ORIGIN

Query Match 71.1%; Score 421; DB 13; Length 695;  
Best Local Similarity 87.2%; Pred. No. 5.7e-112;  
Matches 496; Conservative 0; Mismatches 70; Indels 3; Gaps 3;

QY 23 GAAAGGACATTAACATCAACTATGTATTAAGCTCTGCTGGAAGAAACATATTAAG 82  
DB 695 GGAAGGACATTAACATCAACTATGTATTAAGCTCTGCTGGAAGAAACATATTAAG 636  
QY 83 AACATATGCTTTCTCTATTAAGAAAGCTAGTCTGTAACAAAGCTACTTACC 142  
DB 635 CACAAATATGCTTTCTCTCAAGAAAGCTGCTGCTGTAAGAGAGATTATCACTACC 576  
QY 143 ACAGTATCAGGCTGGAAGACGAGATGTGTTTCCAGGGAACCTTGTGTGCTGCTCC 202  
DB 575 ATGACTATCAGATGTGAGACGAGATGTGTTTCCAGGGAACCTTGTGTGCTGCTCC 516  
QY 203 AGTCACCTTACA-CGCTGTCAAGACCTTGTGATTTGCTCCCTGACACACACCTGAG 261  
DB 515 AATCTCCCAATCTCTGTCGAAAGACTTGTGATTAATCCCTGACACACACCTGAG 456  
QY 262 ACATCCGTTAGAGATGTGATGAGCTGCT-GATGTCTACACAGATGTGAAAGCTGCTG 320  
DB 455 ACATCCGTTAGAGATGTGATGAGCTGCTGATGAGCTGCTACAGGAGTGAACACGCTG 396  
QY 321 GAATGACAGAAATTCATTTTCATGAGGCTCAATGTGCTGAGCTGAGCTGCCAA 380  
DB 395 GAGGCGGAGAAATTCATTTTCATGAGGCTCAATGTGCTGAGCTGAGCTGCCAA 336  
QY 381 GAAGGCTGGAAGACATCCGCTGAGACGAGACCCCAAGTTGTTGCTGATCGGGA 440  
DB 335 GAGGCTGGAAGACATCCGCTGAGACGAGACCCCAAGTTGTTGCTGATCGGGA 276  
QY 441 CCAAGAGACACCAAGCTCAAGAAAGACCAAACTGCGCTTATGACAGATGCTGTTAG 500  
DB 275 CCAAGAGACACCAAGCTCAAGAAAGACCAAACTGCGCTTATGACAGATGCTGTTAG 216  
QY 501 AGGACAAATATTTGTAACCTGTGTGCTGCTCAATCA-AACCTCGCTTGAATTCAGAA 559  
DB 215 AGGACAAATATTTGTAACCTGTGTGCTGCTCAATCAAGTGAACAGTGTTTTGAATTCAGAA 156  
QY 560 AACCTTACAGGTTTCTGTAATGAAGGCC 588  
DB 155 AACCTTACAGGTTTCTGTAATGAAGGCC 127

RESULT 8  
AW543356/c 642 bp mRNA linear EST 31-AUG-2000  
LOCUS C0168H07-3 NIA Mouse E7.5 Extraembryonic Portion cDNA Library Mus  
DEFINITION musculus cDNA clone C0168H07 3', mRNA sequence.  
ACCESSION AW543356  
VERSION AW543356.1 GI:7185773  
KEYWORDS EST.  
SOURCE house mouse.

ORGANISM	Mus musculus																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
----------	--------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

```

Qy 1 GAGAGCTAAACGAAATTCAGAAAAGCATACATACATATGATAGCTCTGC 60
Db 221 GAGAGCTGAATGAAATTCACGAAAGACATATACATATGATAGCTCTCGA 280
Qy 61 CTTGGAGAAACACATATTAAGAACAGTATGCTTTCTTATTAAGAAAGCTATGCT 120
Db 281 CTTGGAGAAACACATATTAAGAACAGTATGCTTTCTTATTAAGAAAGCTATGCT 340
Qy 121 GTAAACAAAGCTACCTCTACACAGCATATCAGGCTGAGAGCGAGATGTTTCCAG 180
Db 341 GTGAAGCAAAATATCCCTACCATGACTATCAGATGAGACACAGACGTTTTCAGG 400
Qy 181 GAACCTTTTGTGCTGTTTCACTCAACCTTACACCGCTGTCAAGAGACTTGTATGTC 240
Db 401 GAGCCCTTGTGTTGTTGTTTCCATTCCTCCCTTATCTGCTGTCAAGAGACTTGTATGTC 460
Qy 241 CCCCTGACACACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGATGTTAC 300
Db 461 CCCTTGACACAACTCCCGAGACTCCGTTAAAGAGATGAGTGTGATGTAC 520
Qy 301 ACAGATGTGAACGTGCTGGAATGACAGATTTCTTTTCAATGGTGACTTCAATGCT 360
Db 521 ACCGATGTGAAGCCAGTGAAGACAGAGATTTCTTCAATGGTGAATTTCAACGCC 580
Qy 361 GGCTGAGCTACGTCCCAAGAAAGCCTGGAAGACATCCGCTGA-GAGACGACCCCA 419
Db 581 GGCTGAGCTATGTCCCAAGAAAGCCTGGAAGACATTTGTTGAGAGACGACCCCA 610
Qy 420 GTTCG-TTTCGCTGATGGGACCAAGAGACACACGCTCAAGAGACAACTGCG 478
Db 641 GTTGTGTTTGGCTGATGGGACCAAGAGACACACGCTCAAGAGAGTACCGCTGTG 700
Qy 479 CCTATGACAGATTCGCTTATAGAGACAAATATTTCACTGTGTGCTTCAATCA 538
Db 701 CCTATGACAGATTCGCTTATAGAGACAAATATTTCACTGTGTGCTTCAATCA 760
Qy 539 ACCCTGCTTTGATTTCCAGAAAGCTTACAGTGTGCTGATCGAAGGCC 588
Db 761 GTGGCGCTTTGACTTTCAGAAAGCTTATGACTGTGTGACGACGAGCCC 810

```

```

RESULT 10      825 bp      mRNA      linear      EST 03-APR-2001
BG530136      602558626f1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:469638 5',
LOCUS          mRNA sequence.
DEFINITION     BG530136
ACCESSION      BG530136
VERSION        BG530136.1 GI:13521673
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 825)
                NIH-MGC http://mgi.nci.nih.gov/.
                National Institutes of Health, Mammalian Gene Collection (MGC)
                Unpublished (1999)
                Contact: Robert Strausberg, Ph.D.
                Email: cgabbs-r@mail.nih.gov
                Tissue Procurement: ATCC
                CDNA Library Preparation: CLONTECH Laboratories, Inc.
                CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
                DNA Sequencing by: Incyte Genomics, Inc.
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LINL at:
                http://image.llnl.gov
                Place: LNCM1524 row: c column: 07
                High quality sequence stop: 743.
                Location/Qualifiers

```

## FEATURES

```

source
1..825
/organism="Homo sapiens"
/db_xref="taxon:9606"

```

```

/clone="IMAGE:469638"
/clone_lib="NIH_MGC_61"
/tissue_type="embryonal carcinoma"
/lab_host="DH10B (TI phage-resistant)"
/notes="Organ: testis; Vector: pNR-LIB (Clontech); Site: 1;
Site1 (ggccgctcggc); Site 2: Site1 (ggccatctggc);
Double-stranded cDNA was prepared from cell line RN. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-ATTCTAGAGCGCCGAGCGCCGACATG-dt(30)-BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."

```

```

BASE COUNT      225 a      204 c      217 g      179 t
ORIGIN
Query Match      67.0%; Score 396.4; DB 12; Length 825;
Best Local Similarity 82.5%; Pred. No. 9.5e-105;
Matches 490; Conservative 0; Mismatches 101; Indels 3; Gaps 3;

```

```

Qy 1 GAGAGCTAAACGAAATTCAGAAAAGCATACATACATATGATAGCTCTGC 60
Db 232 GAGAGCTGAACGAAATTCAGAAAGACATACATATGATAGCTCTGC 291
Qy 61 CTTGGAGAAACACATATTAAGAACAGTATGCTTTCTTATTAAGAAAGCTATGCT 120
Db 292 CTTGGAGAAACACATATTAAGAACAGTATGCTTTCTTATTAAGAAAGCTATGCT 351
Qy 121 GTAAACAAAGCTACCTCTACACAGCATATCAGGCTG-AGACGAGATGTTTCCAG 179
Db 352 GTAAAGAGAGTATATACATACATGACTATCAGATGCGACAGATGTTTCCAG 411
Qy 180 GGAACCTTTGTGCTGTTCCAGTACCTTACACCGCTGTCAAGAGACTTGTGATGT 239
Db 412 GGAACCTTTGTGCTGTTCCAGTACCTTCCCACTGCTGTCAAGAGACTTGTGATGT 471
Qy 240 CCCCTGACACACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGCTGATGTCTA 299
Db 472 CCCCTGACACACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGCTGATGTCTA 531
Qy 300 CACAGATGTGAACGTCGCTGGAATGACAGAAATTTCAATGCTGATGCTCAATGC 359
Db 532 CACAGATGTGAACGTCGCTGGAATGACAGAAATTTCAATGCTGATGCTCAATGC 591
Qy 360 TGCTGACGTAAGTCCCAAGAGGCTGGA-AGACATCCGCTGAGAGCGACCCCA 418
Db 592 CGCTGACGTTATAGTCCCAAGAGGCTGGAAGCAACATCCGCTGAGAGCGACCCCA 651
Qy 419 AGTTGCTTTGCTGATGGGACCAAGAGACACACGCTCAAGAGACCAACTGGC 478
Db 652 GGTTCCTTGTGCTGATGGGACCAAGAGACACACGCTCAAGAGACCAACTGGC 711
Qy 479 CCTATGACAGAT-TGCTGCTTGAAGCAAAATATTGTCACTGCTGTGCTCTCAATCA 537
Db 712 CATATGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 771
Qy 538 AACCTGCTTTGATTTCCAGAAAGCTTACAGTGTGCTGAATCGAAGGCCCTTG 591
Db 772 AACCAAGGCTTTGACTTTCAGAAAGCTTCCAGGCTGACTGTGAAGAGGCCCTTG 825

```

```

RESULT 11      620 bp      mRNA      linear      EST 31-AUG-2000
AW543802      C0174C07-3 N1A Mouse E7.5 Extraembryonic Portion cDNA Library Mus
LOCUS          AW543802/c
DEFINITION     C0174C07-3 N1A Mouse E7.5 Extraembryonic Portion cDNA Library Mus
ACCESSION      AW543802
VERSION        AW543802.1 GI:7186219
KEYWORDS       EST.
SOURCE         house mouse.

```

ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 620)  
AUTHORS Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grahovac  
M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H., Wood,W.H.  
III, Becker,K.G. and Ko,M.S.H.  
TITLE Genome-wide expression profiling of mid-gestation placenta and  
embryo using a 15,000 mouse developmental cDNA microarray  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)  
MEDLINE 20381348  
COMMENT Contact: George J. Kargul  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: cdna@lgsun.grc.nia.nih.gov  
Plate: C0174 row: C column: 07  
Seq primer: -21M13 Forward  
High quality sequence stop: 620  
POLYA=Yes.  
FEATURES Location/Qualifiers  
source 1..620  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="niaEST:C0174C07-3"  
/db\_xref="taxon:10090"  
/clone="C0174C07"  
/clone\_lib="NIA Mouse E7.5 Extraembryonic Portion cDNA  
Library"  
/sex="unknown"  
/dev\_stage="7.5dpc Embryo"  
/lab\_host="DH10B"  
/note="Vector: pSPORT1 (Gibco/BRL Life Technology);  
Site 1: SalI; Site 2: NotI; Total RNAs were extracted from  
5 EFC. The double-stranded cDNA was synthesized by  
Gibco's kit with an Oligo(dT) primer [NotI primer-adaptor  
from Gibco/BRL]  
[5'-pGACTAGTCTAGATCGGCGGCCCTTTTCTTTT-3']  
from 0.8ug of mRNA. The double-stranded cDNAs were  
treated with T4 DNA polymerase and purified by  
ethanol-precipitation. The cDNAs were ligated to  
Lone-linker L1-Sal3 (include SalI sequence). The cDNAs  
were purified by phenol/chloroform and separated from  
free linkers by Centricon 100. Then, cDNAs were amplified  
by long-range high fidelity PCR using Takara's Ex Taq  
polymerase. Then, the cDNAs were purified by  
phenol/chloroform and by Centricon 100. The cDNAs were  
digested with SalI and NotI enzymes. Then, the cDNAs were  
size selected by Gibco's Size Fractionation Column. The  
cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid  
vector. The DH10B E. coli host was transformed with the  
ligation mixture by chemical method. The library was  
constructed by Xiaohong Wang."  
BASE COUNT 160 a 159 c 143 g 158 t  
ORIGIN  
Query Match 66.1%; Score 391.2; DB 10; Length 620;  
Best Local Similarity 82.8%; Pred. No. 2.8e-103;  
Matches 447; Conservative 0; Mismatches 93; Indels 0; Gaps 0;  
QY 53 GCTCTGCCTTGAAGAAACACATATAAAGACAGATGCTCTTCTCTATAAAGAAAGC 112  
DB 620 GTTCTCGATGTGAAGAAACACGATACAGAGAGAGATGCTCTCTACAGGAGAC 561  
QY 113 TAGTGTCTGTAAACAAAGTACTCTTACCACGACTATCAGGCTGGAGACGAGATGTT 172  
DB 560 TGTGTCTGTGAGACAAATACCATCCTACCTATCAGGATGGAGACAGACGCTGT 501  
QY 173 TTTCCAGGGAACCTTTGTGGTCTGGTCCAGTCCACCTACCGGTGTCAAGACTTCG 232  
DB 500 TTTCCAGGGAACCTTTGTGGTCTGGTCCATTTCCCTTTACTGTGTCAAGACTTCG 441  
QY 233 TGATTGTCCCTGCACACCCCTGAGACATCCCGTTAGAGAGATTGTAGCTGGCTG 292

Db 440 TGATTGTCCCTTGCACACAACTCCGAGACTCGTTAAAGAGATAGATGCTGGTCG 381  
QY 293 ATGTCTACACAGATGTGAACCGTGCCTGGAATGAGAGAAATTTCAATTTTATGGTGACT 352  
Db 380 ATGTCTACACGGATGTGAGAGCCAGTGTGAAGACAGAGAAATTTCAATTTTATGGTGACT 321  
QY 353 TCAATGTGGTGGAGCTACGTCCTCCCAAGAGAGGCTGGAAGACATCCGCTGAGGACGG 412  
Db 320 TCAACGCCGGCTGTAGCTATGTCTCCCAAGAGAGGCTGCGAGAACATTCGTTTGAAGACGG 261  
QY 413 ACCCAAGTTCTGTTGGCTGATCGGGGACCAAGAGACACACCGTCAAGAGAGACAA 472  
Db 260 ACCCAAGTTGTTGGCTGATCGGGGACCAAGAGACACATACGTTCAAGAGAGTACCA 201  
QY 473 ACTGGCCCTATCAGAGATCGTGTAGAGGACAAAATTTGTCAACTCTGTTGCTCCTC 532  
Db 200 GCTGTGCTATGACAGGATGTGCTTTGTGGACAAGAGATAGTCAACTCCGTTGACCCC 141  
QY 533 AATCAACCTCGCTTGTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592  
Db 140 GTTCCAGTGGCGTCTTGTGACTTTTCAGAAAGCTTATGACTTGTCTGAGGAGGAGGCTGG 81  
RESULT 12  
BI460950 603207187F1 NIH\_MGC\_97 Homo sapiens cDNA clone linear EST 21-AUG-2001  
LOCUS mRNA sequence.  
DEFINITION BI460950.1 GI:15251606  
ACCESSION BI460950  
VERSION BI460950  
KEYWORDS EST  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 628)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs@mail.nih.gov  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM11688 row: n column: 22  
High quality sequence stop: 628.  
FEATURES Location/Qualifiers  
source 1..628  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5272917"  
/clone\_lib="NIH\_MGC\_97"  
/lab\_host="DH10B"  
/note="Organ: testis; Vector: pBluescriptR (modified  
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgaag  
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',  
size-selected for average insert size 2.2 kb and  
normalized to ROT 5. This is a primary library enriched  
for full-length clones and constructed using the  
cap-trapper method (Carninci, in preparation). Library  
constructed by M. Brownstein (NIH/NHGRI, National  
Institutes of Health). Note: this is a NIH\_MGC Library."  
BASE COUNT 181 a 144 c 166 g 137 t  
ORIGIN  
Query Match 65.8%; Score 389.4; DB 13; Length 628;  
Best Local Similarity 90.1%; Pred. No. 9.4e-103;



	Matches	417;	Conservative	0;	Mismatches	46;	Indels	0;	Gaps	0;	
Qy	1	GAGAGCTTAAACG	AAATTCAG	AAAAAGGCATTA	CACTTA	CAACTATG	ATTAGCTTCG	CC	60		
Db	166	GAGAGCTGTAAG	AAATTCAG	AAAGGCA	TAACAT	AACTATG	ATTAGCTTCG		225		
Qy	61	CTTGAGAGAA	CACATAT	AAAGAC	AGTATG	CTTCTCTAT	ATAAGAAAGCTAG	GTCT	120		
Db	226	CTTGAGAGAA	CACATAT	AAAGAC	ATATATG	CTTCTCTAC	AAAGAAAGCTAG	GTCT	285		
Qy	121	GTAAACA	AAAGTAC	CTTA	CCACCA	CTATCA	GCGTGGAGAGCGCAGATGT	TTTTCCAG	180		
Db	286	GTAAGAGAG	GTATATCA	CTACATG	CTATCA	GAGAGAGAGACG	CAATATGT	TTTTCCAG	345		
Qy	181	GAACCTTTGT	TGTGTG	TTCCAGTAC	CCCTAC	CCGCGTCA	AGGACTCGT	GTATTC	240		
Db	346	GAGCCCTTTG	TGTGTG	TTCCATCT	CCCACT	CGTGTCA	AAAGACTCGT	GTATTC	405		
Qy	241	CCCCCTG	CACACC	CCCCCTG	AGACAT	CCGTTAG	AGATGATG	AGCTGCTAGTGT	CTAC	300	
Db	406	CCCCGCA	CACCAC	CCCCAG	AGACAT	CCGTTAG	AGATGATG	AGCTGCTAGTGT	CTAC	465	
Qy	301	ACAGATG	GAACCGT	CGCTG	GAATG	GAAGATTA	TCATTTAT	GATGGGTACT	TCATGT	360	
Db	466	ACGACG	TGAACAC	CCGCTG	GAAGCG	GAATTTCA	TTTATG	GGTACTTCA	TGCT	525	
Qy	361	GGCTG	CAGCTAC	GTCC	CCAA	AGGCTT	GGAAGAC	ATCCGCTT	GAGACG	420	
Db	526	GGCTG	CAGCTAC	GTCC	CCAA	AGGCTT	GGAAGAC	ATCCGCTT	GAGACG	585	
Qy	421	TTTCG	TTTGGCT	GATCG	GGGAC	CAAGAG	CACAC	CGGTCA	AA	463	
Db	586	TTTGT	TTGGCT	GATCG	GGGAC	CAAGAG	CACAC	CGGTCA	AA	628	

ACCESSION AW915564  
VERSION AW915564.1 GI:8081263  
KEYWORDS EST.  
SOURCE Rattus sp.  
ORGANISM Rattus sp.  
REFERENCE 1 (bases 1 to 668)  
AUTHORS Lee, N.H., Glodek, A., Chandra, I., Mason, T.M., Quackenbush, J.,  
Kerlavage, A.R., and Adams, M.D.  
TITLE Rat Genome Project: Generation of a Rat EST (EST) Catalog & Rat  
Gene Index  
JOURNAL Unpublished (1998)  
COMMENT Contact: Lee, NH  
The Institute for Genomic Research  
9712, Medical Center Drive, Rockville, MD 20850, USA  
Tel: (301)-838-3529  
Fax: (301)-838-0208  
Email: nhlee@tigr.org  
This clone is available through the ATCC, contact the ATCC  
tel#703-365-2700 for further information  
Seq primer: M13 Reverse.  
FEATURES  
source  
1..668  
/organism="Rattus sp."  
/db xref="taxon:10118"  
/clone="RGICW80"  
/dev\_stage="embryo 8, 12, 18 dpc"  
/note="vector: p773Pac; Site\_1: EcoRI; Site\_2: NotI"  
BASE COUNT 172 a 160 c 168 g 168 t  
ORIGIN  
Query Match 64.8%; Score 383.8; DB 10; Length 668;  
Best Local Similarity 85.6%; Pred. No. 4.2e-101;  
Matches 439; Conservative 0; Mismatches 72; Indels 2; Gaps 1;  
QY 80 AAGAACAGTATGCTTTCTCTATAAGAAAGCTAGTGTCTGTAAACAAAGCTACCTCT 139  
DB 2 AAGAACAGTATGCTTTCTCTATAAGAAAGCTAGTGTCTGTGTGAAGCAAA--ACCTCT 59  
QY 140 ACCAGACTATCAGCTGGAGCGAGATGTGTTTTCCAGGAAACCTTTGTGGTCTGCT 199  
DB 60 ACCATGACTATCAGGATGAGACAGACGCTGTTTCCAGGAGCCCTTTGTGGTCTGCT 119  
QY 200 TCCAGTCACTTACACCGCTGTCAAGGACTTCGTGATTTGTCCTCCCTGCACACCACTCCCTG 259  
DB 120 TCCAGGCGCCCTTCACTGTCGCAAGGACTTCGTGATTTGTCCTCCCTGCACACCACTCCCTG 179  
QY 260 AGACATCCGTTAGAGAGATTGATGCTGGCTGATGTCTACACAGATGTGAACGCTCGCT 319  
DB 180 AAACCTCCGTTAAAGAGATGATGATGCTGGCTGACGCTCTACACGATGTGAGAACGAT 239  
QY 320 GAATGACAGAGATTTCATTTTCATGGGTGATCTCAATGCTGCTGCTGCTGCTGCTGCTGCT 379  
DB 240 GGAAGGACAGAGATTTCATTTTCATGGGTGATTTCAATGCTGGCTGCTGCTGCTGCTGCTGCT 299  
QY 380 AGAAGGCTTGAAGGACATCCCGCTGAGGACGAGACCCCAAGTTTCGTTTGGCTGATCGGG 439  
DB 300 AGAAGGCTTGAAGACATCCGTTTGAAGACAGACACCCCACTTTGTTGGCTGATTTGGG 359  
QY 440 ACCAAGAGACACCGCTTGAAGAGAGACAAATCGCCCTTATGACAGGATCGTCTTA 499  
DB 360 ACCAAGAGACACCGCTTGAAGAGAGACACCAAGCTGTGCTTATGACAGGATTTGCTTC 419  
QY 500 GAGGACAAATATTTGCAACTCTGTTGCTTCAATCAAACTCGTCTTTGATTTTCCAGA 559  
DB 420 GCGGCAAGAGATGATGCAACTCTGTTGCTTCCCGCTTCAGTGGCGCTTTTGAATTTTCA 479  
QY 560 AAGCTTACAGGTTGCTGTAATCGAAGGCCCTGG 592  
DB 480 AAGCTTATGATTTGCTGGAAGAGAGAGGCCCTGG 512

RESULT 15  
AI927844/c  
LOCUS  
DEFINITION  
AI927844.1  
GI:5663808  
EST.  
PRECUSOR ; mRNA Sequence.  
AI927844  
AI927844.1  
GI:5663808  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 784)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html  
Insert Length: 739 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 426.

FEATURES  
source  
1..784  
/organism="Homo sapiens"  
/db xref="taxon:9606"  
/clone="IMAGE:2462752"  
/clone\_lib="NCI CGAP\_Kid11"  
/lab\_host="DH10B"  
/note="Organ: Kidney; Vector: p773D-Pac (Pharmacia) with  
a modified polylinker; Site\_1: Not I; Site\_2: Eco RI;  
Plasmid DNA from the normalized library NCI CGAP Kid3 was  
prepared, and ss circles were made in vitro. Following HAP  
purification, this DNA was used as tracer in a subtractive  
hybridization reaction. The driver was PCR-amplified cDNAs  
from a pool of 5,000 clones made from the same library  
(cloneids 1322376-1323911, 1456007-1456775, and  
1500552-1502855). Subtraction by Bento Soares and M.  
Fatima Bonaldo."

BASE COUNT 173 a 188 c 179 g 242 t 2 others  
ORIGIN  
Query Match 64.8%; Score 383.6; DB 9; Length 784;  
Best Local Similarity 86.7%; Pred. No. 5.1e-101;  
Matches 422; Conservative 0; Mismatches 65; Indels 0; Gaps 0;  
QY 106 GAAAGCTAGTCTCTGTAACAAAGCTACCTCTACACGACTATCAGGCTGGAGACGCA 165  
DB 667 GAAAGCTGGTCTGTGTAAGAGAGATTATCATTACCATGATCATGAGTGGAGACGCA 608  
QY 166 GATGTTTTCAGGGAACCTTTTGTGTCTGTGTTCCAGTCCACCTACACCGCTGTCAAG 225  
DB 607 GATGTTTTCAGGAGCCCTGTGTGTCTGTGTTCCAACTCCCCACACTGCTGCAAA 548  
QY 226 GACTTCGATGTTCCCTCTGCACACCACTCCCTGAGACATCCGTTAGAGATGATGAG 285  
DB 547 GACTTCGATGTTATCCCTCTGCACACCACTCCCTGAGACATCCGTTAGGATCGATGAG 488  
QY 286 CTGGCTGATGCTACACAGATGTGAACGTCGCTGGAGTGCAGAGATTTTCATTTTCATG 345  
DB 487 TTGGTTGAGTCTACACGACGTGAACACCCCTGGAGCGGAGATTTTCATTTTCATG 428

```

Oy 346 GGTGACTTCAATGCTGGCTGCAGCTACGTCCCAAGAGGCTTGAAGAGACATCCGCTG 405
    |||||
Db 427 GGTGACTTCAATGCTGGCTGCAGCTACGTCCCAAGAGGCTTGAAGAGACATCCGCTG 368
    |||||
Oy 406 AGGACCGAGCCCCAAGTTCGTTGGCTGATCGGGACCAAGAGACACGTCAGAGAG 465
    |||||
Db 367 AGGACTGACCCCGAGTTGTTGGCTGATCGGGACCAAGAGACACGTCAGAGAG 308
    |||||
Oy 466 AGCACAACCTGCGCTATGACAGGATCGTCTTAGAGACAAAATATGTCATCTGTG 525
    |||||
Db 307 AGCACCACACTGTGCAATGACAGGATTTGCTTAGAGACAAATGTCAGTTCTGTT 248
    |||||
Oy 526 GGTCTCAATCAAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGTTGTCGAATCGAG 585
    |||||
Db 247 GTTCCCAAGTCAAACAGTGTGTTTGAAGTTCCAGAAAGCTTACAAAGTGAAGAGAG 188
    |||||
Oy 586 GCCCTGG 592
    |||||
Db 187 GCCCTGG 181
    |||||

```

Search completed: December 9, 2002, 12:18:03  
 Job time : 1339.4 secs

**This Page Blank (uspto)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 08:59:53 ; Search time 36.0918 Seconds  
(without alignments)  
5030.295 Million cell updates/sec

Title: US-09-905-114-1

Perfect score: 592  
Sequence: 1 gagagcgaacgaagcattc.....gtcgcacgaagcctcg 592

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/pdata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/pdata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/pdata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/pdata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/pdata/1/ina/PCUS.COMB.seq:\*  
6: /cgn2\_6/pdata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	452.8	76.5	1208	1	US-08-640-765A-2 Sequence 2, Appl
2	452.8	76.5	1208	3	US-09-073-613-2 Sequence 2, Appl
3	149.4	25.2	1039	1	US-07-895-300A-2 Sequence 2, Appl
4	149.4	25.2	1039	4	US-09-638-112-2 Sequence 2, Appl
5	149.4	25.2	1039	5	PCT-US93-05136-2 Sequence 2, Appl
6	149.4	25.2	1039	5	PCT-US93-10519-1 Sequence 1, Appl
7	147.8	25.0	1039	1	US-08-458-367-2 Sequence 2, Appl
8	117.2	19.8	1615	2	US-08-491-988-15 Sequence 1, Appl
9	117.2	19.8	1624	2	US-08-491-988-14 Sequence 1, Appl
10	117.2	19.8	1648	2	US-08-491-988-13 Sequence 1, Appl
11	115.2	19.5	1055	2	US-08-468-012A-1 Sequence 1, Appl
12	115.2	19.5	1055	4	US-09-054-989-1 Sequence 1, Appl
13	35.2	5.9	1427	4	US-09-173-581-10 Sequence 10, Appl
14	35.2	5.9	1427	4	US-09-420-915-10 Sequence 10, Appl
15	32.8	5.5	1293	2	US-08-924-440-1 Sequence 1, Appl
16	32.2	5.4	289	4	US-09-007-005-17 Sequence 17, Appl
17	32.2	5.4	289	4	US-09-244-796-17 Sequence 17, Appl
18	32.2	5.4	7218	1	US-08-232-463-14 Sequence 14, Appl
19	31.8	5.4	1511	2	US-08-809-763-3 Sequence 3, Appl
20	31.6	5.3	1001	3	US-08-956-253-3 Sequence 3, Appl
21	31.6	5.3	1001	4	US-09-641-638-90 Sequence 90, Appl
22	31.2	5.3	248	3	US-08-513-974B-300 Sequence 300, Appl
23	31.2	5.3	248	3	US-08-513-974B-301 Sequence 301, Appl
24	30.6	5.2	1920	1	US-08-487-753-1 Sequence 1, Appl
25	30.6	5.2	1920	2	US-08-480-065-1 Sequence 1, Appl
26	30.6	5.2	1920	3	US-08-487-744-1 Sequence 1, Appl
27	30.6	5.2	1920	5	PCT-US93-09167-1 Sequence 1, Appl

C 28	30.6	5.2	2900	2	US-09-027-337-9 Sequence 9, Appl
C 29	30.6	5.2	2900	4	US-09-644-600-9 Sequence 9, Appl
C 30	30.6	5.2	3147	2	US-09-027-337-1 Sequence 1, Appl
C 31	30.6	5.2	3147	4	US-09-644-600-1 Sequence 1, Appl
C 32	30.6	5.2	3147	4	US-09-644-600-18 Sequence 18, Appl
C 33	30.6	5.1	1734	4	US-09-484-970B-63 Sequence 63, Appl
C 34	30.6	5.1	2071	1	US-07-923-724-1 Sequence 1, Appl
C 35	30.6	5.1	2071	2	US-08-609-426A-1 Sequence 1, Appl
C 36	30.6	5.1	2071	2	US-08-374-652C-3 Sequence 3, Appl
C 37	30.6	5.1	4049	1	US-08-162-809-17 Sequence 17, Appl
C 38	30.6	5.1	4097	1	US-08-162-809-11 Sequence 11, Appl
C 39	29.6	5.0	352	4	US-09-228-986-21 Sequence 21, Appl
C 40	29.6	5.0	560	3	US-09-042-426-5 Sequence 5, Appl
C 41	29.6	5.0	560	4	US-09-291-238-5 Sequence 5, Appl
C 42	29.6	5.0	560	4	US-09-330-760-5 Sequence 5, Appl
C 43	29.6	5.0	560	4	US-09-328-473-5 Sequence 5, Appl
C 44	29.6	5.0	560	4	US-09-330-727-5 Sequence 5, Appl
C 45	29.6	5.0	560	4	US-09-329-169-5 Sequence 5, Appl

## ALIGNMENTS

RESULT 1  
US-08-640-765A-2  
Sequence 2, Application US/08640765A  
Patent No. 5821103  
GENERAL INFORMATION:  
APPLICANT: Tanuma, Sei-ichi  
TITLE OF INVENTION: NOVEL DEOXYRIBONUCLEASE  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESSES:  
ADDRESS: Leydig, Volt & Mayer  
STREET: Two Prudential Plaza, Suite 4900  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/640,765A  
FILING DATE: 06-MAY-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6239518  
FILING DATE: 06-SEP-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Green, Robert F.  
REGISTRATION NUMBER: 27555  
REFERENCE/DOCKET NUMBER: 73362  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 616-5600  
TELEFAX: (312) 616-5700  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1208 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 12..941  
US-08-640-765A-2  
Query Match 76.5%; Score 452.8; DB 1; Length 1208;  
Best Local Similarity 85.3%; Pred. No. 5,4e-132;  
Matches 505; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

REGISTRATION NUMBER: 35243  
REFERENCE/DOCKET NUMBER: 85519  
TELECOMMUNICATION INFORMATION:

Qy	1	GAGAGCTAAACCGAAATTCAAGAAAGGCATAACAATATGTGTGATTAGCTCTCGC	60
Db	243	GAGAAGCTGAATCGAAACTCAGAAAGAGCAGACATACGTGATAGCTCTGG	302
Qy	61	CTTTGGAGAAACACATATAAAGAACAGTATGCCCTTTCTCTATAAAGAAAGCTAGTGCT	120
Db	303	CTTTGGAAAGAAACACATATAAAGAACAGTATGCCCTTCTACAGAGAGAAGCTGGTGCT	362
Qy	121	GTAAAACAAAAGCTACCTCTACACGACTATCAGGCTGGAGACGAGATGTGTTTTCCAGG	180
Db	363	GTGAGGCAAAATACCTCTACCATGACTATCAGGATGGAGACACAGACGTGTTTTCCAGG	422
Qy	181	GAACCCCTTTGTGGTCTGGTTCCAGTCAACCTACACCGCTGTCAAGACATTCGTGATTGTC	240
Db	423	GAGCCCTTTGTGGTTTGGTTTCCAGGCGCCCTTCACCTGCTGCCAAGACATTCGTGATTGTC	482
Qy	241	CCCCGTGCACACCCCTGAGAGACATCCGTTAGAGAGATGATGAGCTGGCTGATGTCATC	300
Db	483	CCCTTGACACAACTCCTGAAACCCTCGTTTAAAGAGATAGATGAGCTGGCTGACGTCAC	542
Qy	301	ACAGATGTGAAACGTCCTGGAAATCGAGAGAAATTCATTTTCATGGGTGACTTCAATGCT	360
Db	543	ACGATGTTAGAGACATGGAAGGCAGAGATTTTCATCTTCATGGTGATTTCAATGCT	602
Qy	361	GGCTGCAGCTACGTCCTCCAAAGAGCCTCGAAGACATCCGCTTGAGGACGAGCCCCAAG	420
Db	603	GGCTGCAGCTACGTCCTCCAAAGAGCCTCGAAGAAATCCGTTTGAGGACAGACCCCCAAC	662
Qy	421	TTGCTTTTGGCTGATCGGGGACCAAGAGACACACCGTCAAGAGAGGACAAACTCGGCC	480
Db	663	TTTGTGTTGGCTGATTGGGGACCAAGAGGACACCAACGCTCAAGAGAGACACAGCTGTGCC	722
Qy	481	TATCAGAGGATCGTGCTTTAGAGGACAAATATGTCACATCTCGTGGTCCCTCAATCAAAAC	540
Db	723	TATCAGAGGATGTGCTTTCGGCGAACAGAGATGATCAATCTGTGTTTCCCGCTCCAGT	782
Qy	541	CTCGCTTTTGATTTCCAGAAAGCTTACAGGTTGTCTCAATTCGAAGGCCCTTGG	592
Db	783	CGCGCTTTGATTTTCAGAAAGCTTATCAGTTGTCTCAAGAGAGAGGCCCTTGG	834

```

Db      543  ACGGATGTTAGAACGACGATGGAAGCAGAGATTTTCATCTTTCATGGGTGATTTCAATGCT   602
QY      361  GGCTGCAGCTACGTCCCCCAAGAGCCCTGGGAAGAGACATCCGCCTGAGACAGCACCCCAAG   420
Db      603  GGCTGCAGCTACGTCCCCCAAGAGCCCTGGGAAGAACATCCGTTTGAGGACAGACCCCAAC   662
QY      421  TTCGTTTGGCTGATCGGGGACCAAGAGAGACACACCGGTCAAGAAGAGACACAAACTGCGCC   480
Db      663  TTTGTTTGGCTGATTTGGGGACCAAGAGAGACACACCGGTCAAGAAGAGACACAGCTGTGCC   722
QY      481  TATGACAGGATGTCGCTTAGAGGACAAATATTTGTCAACTCTGGTGGTGCTCTCAATCAAC   540
Db      723  TATGACAGGATTTGCTGTTCCGGGACAAAGAGATAGTCAACTCTGTGTGTTCCCGCTCCAGT   782
QY      541  CTCGTCTTTGATTTCACAAAGCATTACAGGTGTCCTGAATCGAAGGCCCTGG   592
Db      783  GCGCTCTTTGACTTTCAGAAAGCTTATCAGTTGTCTGAAGAGAGGCCCTGG   834

RESULT 3
US-07-895-300A-2
; Sequence 2, Application US/07895300A
; Patent No. 5279823
; GENERAL INFORMATION:
; APPLICANT: Prenz, John
; APPLICANT: Shire, Steven J.
; APPLICANT: Sliwowski, Mary B.
; TITLE OF INVENTION: PURIFIED FORMS OF DNase
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
;

```

RESULT 2  
US-09-073-613-2  
; Sequence 2, Application US/09073613  
; Patent No. 6143875  
; GENERAL INFORMATION:  
; APPLICANT: Tanuma, Sei-ichi  
; TITLE OF INVENTION: ANTIBODY TO NOVEL DEOXYRIBONUCLEASE  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Leydig, Voit & Mayer  
; STREET: Two Prudential Plaza, Suite 4900  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/09/073,613  
; FILING DATE: 06-MAY-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/640,765  
; FILING DATE:  
; APPLICATION NUMBER: JP 6239518  
; FILING DATE: 06-SEP-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: LARCHER, Carol

ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/895,300A  
FILING DATE: 19920608  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Hensley, Max D.  
REGISTRATION NUMBER: 27,043  
REFERENCE/DOCKET NUMBER: 747  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1039 bases  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear

US-07-895-300A-2

Query Match 25.2%; Score 149.4; DB 1; Length 1039;  
Best Local Similarity 56.7%; Pred. No. 4,4e-37;  
Matches 317; Conservative 0; Mismatches 236; Indels 6; Gaps 2;

QY 34 ACATACCACTATGTGATGATCTCGCGCTTGGAAGAAACATTAAGAAACAGTATGCC 93  
DB 409 ACCATATACCTAGTGTGATGATGACCCCTGGAGGAGACGCTTAAGAGCGCTACCTG 468  
QY 94 TTTCTATTAAGAAAAAGTAGTGTCTGTAAACAAAGCTACCTCTACCAAGA---CTAT 150  
DB 469 TTGCTGTAGAGGCGCTGACCGAGTGTCTGCGGTGAGACAGCTACTACATGATGCTGC 528  
QY 151 CAGGCTGGAGACCGAGATGTTTTCAGAGGAACCTTTGTGTCTGTTCCAGTCAACC 210  
DB 529 GAGCCCTGCGGGAACGACCTTCACCGAGAGCCACCATTTGTCAAGTTCTTCTCCCG 588  
QY 211 TACACCGCTGTCAAGAGCTCGTATGTTGCCCTGACACACCCCTGAGACATCCGTT 270  
DB 589 TTACACAGAGTCAGGAGGTTTGCATTTGTCCTGATGCGGCGCGGAGACGACGTA 648  
QY 271 AGAGAGATTGATGAGCTGCTGATGTCTACACAGATGTGAACCTCCCTGAAATGACAG 330  
DB 649 GCGGAGATGACCGCTCTCTATGACGTCTACCTGATGCCAAGAAATGGGGTTGGAG 708  
QY 331 AATTTCAATTTTCAATGGGTGACTTCAATGCTGCTGACGCTACGCTCCCAAGAGGCTGG 390  
DB 709 GACGTCTATGTTGATGGCGACTTCMAATGCGGGCTGACGCTATGTGAGAACCTCCCACTGG 768  
QY 391 AAGGACATCCGCGCTGAGAGAGGAGCCCAAGTTGCTTTGGGCTGATCGGGGACCAAGAGAC 450  
DB 769 TCATCTATCCGCGCTGTGAGACAAAGCCCACTTCCAGTGGCTGATCCCGACAGGCTGAC 828  
QY 451 ACCAGGCTCAAGAGAGACAAACCTGCGCTATGACAGATCGTCTTAAGACCAAAAT 510  
DB 829 ACCACAG---CTACACCCAGCACTGAGCTATGACAGATCGTGTGAGAGGATCTG 885  
QY 511 ATTGTCACTCTGTGTCTCTCAATCAAACTCTGTTTGTATTTCCAGAAAGCTTACAGG 570  
DB 886 CTCGAGGCGCGTGTGTTCCGACTCGGCTCTTCCCTTAATCTCCAGGCTGCTATGAGC 945

QY 571 TTGCTGAATCGAAGGCC 589  
DB 946 CTGAGTGACCAACTGGCCC 964

RESULT 4  
US-09-638-112-2  
Sequence 2, Application US/09638112  
Patent No. 6440412

GENERAL INFORMATION:  
APPLICANT: Frenz, John  
APPLICANT: Shire, Steven J.  
APPLICANT: Sliwkowski, Mary B.  
TITLE OF INVENTION: PURIFIED FORMS OF Dnase  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/638,112  
FILING DATE: 09-Aug-2000  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/942561  
FILING DATE: 01-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Johnston, Sean A.  
REGISTRATION NUMBER: 35,910  
REFERENCE/DOCKET NUMBER: P0747C7  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-3562  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1039 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Single  
TOPOLOGY: linear

US-09-638-112-2

Query Match 25.2%; Score 149.4; DB 4; Length 1039;  
Best Local Similarity 56.7%; Pred. No. 4,4e-37;  
Matches 317; Conservative 0; Mismatches 236; Indels 6; Gaps 2;

QY 34 ACATACCACTATGTGATGATGCTCGCGCTTGGAAGAAACATTAAGAAACAGTATGCC 93  
DB 409 ACCATATACCTAGTGTGATGATGACCCCTGGAGGAGACGCTTAAGAGCGCTACCTG 468  
QY 94 TTTCTATTAAGAAAAAGTAGTGTCTGTAAACAAAGCTACCTCTACCAAGA---CTAT 150  
DB 469 TTGCTGTAGAGGCGCTGACCGAGTGTCTGCGGTGAGACAGCTACTACATGATGCTGC 528  
QY 451 CAGGCTGGAGACCGAGATGTTTTCAGAGGAACCTTTGTGTGCTGTTCCAGTCAACC 210  
DB 529 GAGCCCTGCGGGAACGACCTTCACCGAGAGCCACCATTTGTCAAGTTCTTCTCCG 588  
QY 529 GAGCCCTGCGGGAACGACCTTCACCGAGAGCCACCATTTGTCAAGTTCTTCTCCG 588  
DB 529 GAGCCCTGCGGGAACGACCTTCACCGAGAGCCACCATTTGTCAAGTTCTTCTCCG 588  
QY 211 TACACCGCTGTCAAGAGCTCGTATGTTGCCCTGACACACCCCTGAGACATCCGTT 270  
DB 589 TTACACAGAGTCAGGAGGTTTGCATTTGTCCTGATGCGGCGCGGAGACGACGTA 648  
QY 271 AGAGAGATTGATGAGCTGCTGATGTCTACACAGATGTGAACCTGCTGGAATGACAG 330  
DB 649 GCGGAGATGAGAGGCTCTCTATGACGTCTACCTGATGTCCAAAGAAATGGGGCTTGGAG 708





```

; LENGTH: 1039 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
PCT-US93-10519-1

```

Query Match	25.2%	Score 149.4	DB 5	Length 1039
Best Local Similarity	56.7%	Pred. No. 4.4e-37		
Matches 317	Conservative 0	Mismatches 236	Indels 6	Gaps 2

Qy	34	ACATACAACTGTGATAGCTCTGCGCTTGGAGAAACAATTTAAAGAACATATAGCC	93
Db	409	ACCTATCACTACGTGTGTAGTAGACCACTGGAGCGAAACAGCTATTAAAGACCGCTACTCG	466
Qy	94	TTTCTCTATTAAGAAAAGCTAGTGTCTGTATAAACAAAGCTACCTTACCAAGA---CTAT	150
Db	469	TTCTGTATACAGCGCTTGACCAAGTGTCTGGGTGAGACAGCTACTACTACGATGATGGCTGC	528
Qy	151	CAGGCTGGAAGACGAGATGTGTTTTCCAGGGAACCTTTGTGTGTGTGTTCCAATCACCC	210
Db	529	GAGGCTCTGGGGAAACGACACCTTCAACCGAGAGCCAGCCATTTGACAGTCTTCTCCCG	588
Qy	211	TACACCGCTGTCAAGACACTTCGTATTGTCCCCCTGCACACACCCCTGAGACATCCGTT	270
Db	589	TTTCAAGAGGTCAAGGAGATTGGCATTTGTCTCCCTGCATGCGGCGCCGGGGAGCGACGTA	648
Qy	271	AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAACCGTGGCTGGATCGAG	330
Db	649	GCCGAGATCGACGCTCTCTTATGACGTCTACCTGGATGTCCAAGAAATGGGCGTTGGAG	708
Qy	331	AATTTCATTTTTCATGGGTGATTTCAATGTCTGGCTGCAGCTACGTTCGCCAAGAGAGGCTGG	390
Db	709	GACGTCACTGTTGAATGGGCGACTTCAATGCGGCGCTGCAGCTATGTGAAGACCTCCCACTGG	768
Qy	391	AAGGACATCCGCTGAGAGACGGAACCCCAAGTTGCTTTGGCTGATCGGGGACCAAGAGAC	450
Db	769	TCATCTCATTCGCGCTGTGTGACACAGCGCCCACTTTCAGATGGCTGATCTCCCGACAGCGCTGAC	828
Qy	451	ACCAACGGTCAAGAAAGACACAAATCGCGCTTATGACAGATCGTGTCTTAAAGGACAAAT	510
Db	829	ACCAACAG---CTACACCCAGCGCACTGTGGCTATGACAGAAATCGTGTGTGGACGGGATGCTG	885
Qy	511	ATTGTCACTCTGTGTGTCTCTCAATCAAACCTCGTCTTTGATTTTCCAGAAAGCTTACAG	570
Db	886	CTCCGAGGCGCGTGTGTGTCTCCGACTCGGCTCTTCCCTTAACTTCCAGAGCTGCTATGGC	945
Qy	571	TTGTCTGATTCGAAGGCCCC	589
Db	946	CTGAGTGCACATGCGGCC	964

US-PATENT 7  
US-08-458-367-2  
Sequence 2, Application US/08458367  
Patent No. 5783433  
GENERAL INFORMATION:  
APPLICANT: Frenz, John  
APPLICANT: Shire, Steven J.  
APPLICANT: Sliwowski, Mary B.  
TITLE OF INVENTION: PURIFIED FORMS OF DNase  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)





Mon Dec 9 13:45:01 2002

FILING DATE: 5 MAY 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: MULLINS, J.G.  
REGISTRATION NUMBER: 33,073  
REFERENCE/DOCKET NUMBER: 325800-428  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1055 BASE PAIRS  
TYPE: NUCLEIC ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: LINEAR  
MOLECULE TYPE: CDNA  
US-08-468-012A-1

Query Match 19.5%; Score 115.2; DB 2; Length 1055;  
Best Local Similarity 58.6%; Pred. No. 2.2e-26;  
Matches 198; Conservative 1; Mismatches 139; Indels 0; Gaps 0;  
QY 160 GAGCAGATGCTTTTCCAGGAAACCTTTGGTCTGCTTCCAGTCACCCCTACACCGCT 219  
DB 359 GAGGATGACGCTTTTGGCCGGAGCCATTGTTGGCCCACTTCTTTGCCCAGCAATGTC 418  
QY 220 GTCAAGGACTTCGTGATGTTCCACACACACCCCTGCACACACCCCTGCAGATCCGTTAGAGATT 279  
DB 419 CTTCCAGCTGCTGTTGTTCCGCTGCACACACCTCTCTAAGCCGTAGAGAGAGCTG 478  
QY 280 GATGAGCTGGCTGATCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAAATTTTCA 339  
DB 479 AAGCCCTCTACGATGTTCTGAGAGTCTCCAGCACTGGCAGCAAGAGAGCTGATC 538  
QY 340 TTTCATGGGTGACTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 399  
DB 539 CTGCTGGGAGCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 598  
QY 400 CGCCTGAGGACGACCCCAAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 459  
DB 599 GSGCTGCGGACTGAGCAGGCTTCCACTGGTGAITGCGGATGGGAGGAGACACACAGTG 658  
QY 460 AAGAGAGACAACTGCGCTATGACAGATCGTCT 497  
DB 659 CGGCGCAGCACCACCTGACCTATGACCGCTGCTCT 696

RESULT 12  
US-09-054-989-1  
Sequence 1, Application US/09054989  
Patent No. 6251648  
GENERAL INFORMATION:  
APPLICANT: RUBEN, ET AL.  
TITLE OF INVENTION: Human Dnase  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
ADDRESS: CECCHI, STEWART & OLSTEIN  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/054,989  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/468,012

FILING DATE: 6-JUNE-1995  
APPLICATION NUMBER: PCT/US94/04954  
FILING DATE: 5 MAY 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: MULLINS, J.G.  
REGISTRATION NUMBER: 33,073  
REFERENCE/DOCKET NUMBER: 325800-428  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1055 BASE PAIRS  
TYPE: NUCLEIC ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: LINEAR  
MOLECULE TYPE: CDNA  
US-09-054-989-1

Query Match 19.5%; Score 115.2; DB 4; Length 1055;  
Best Local Similarity 58.6%; Pred. No. 2.2e-26;  
Matches 198; Conservative 1; Mismatches 139; Indels 0; Gaps 0;  
QY 160 GAGCAGATGCTTTTCCAGGAAACCTTTGGTCTGCTTCCAGTCACCCCTACACCGCT 219  
DB 359 GAGGATGACGCTTTTGGCCGGAGCCATTGTTGGCCCACTTCTTTGCCCAGCAATGTC 418  
QY 220 GTCAAGGACTTCGTGATGTTCCACACACACCCCTGCACACACCCCTGCAGATCCGTTAGAGATT 279  
DB 419 CTTCCAGCTGCTGTTGTTCCGCTGCACACACCTCTCTAAGCCGTAGAGAGAGCTG 478  
QY 280 GATGAGCTGGCTGATCTACACAGATGTGAAACGTCGCTGGAATGCAGAGAAATTTTCA 339  
DB 479 AAGCCCTCTACGATGTTCTGAGAGTCTCCAGCACTGGCAGCAAGAGAGCTGATC 538  
QY 340 TTTCATGGGTGACTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 399  
DB 539 CTGCTGGGAGCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 598  
QY 400 CGCCTGAGGACGACCCCAAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 459  
DB 599 GSGCTGCGGACTGAGCAGGCTTCCACTGGTGAITGCGGATGGGAGGAGACACACAGTG 658  
QY 460 AAGAGAGACAACTGCGCTATGACAGATCGTCT 497  
DB 659 CGGCGCAGCACCACCTGACCTATGACCGCTGCTCT 696

RESULT 13  
US-09-173-581-10  
Sequence 10, Application US/09173581A  
Patent No. 6013455  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Tang, Y. Tom  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Yue, Henry  
APPLICANT: Guegler, Karl J.  
APPLICANT: Corley, Neil C.  
APPLICANT: Gorgone, Gina  
APPLICANT: Azimzal, Yalda  
APPLICANT: Lu, Aina  
TITLE OF INVENTION: Protein Kinase Homologs  
FILE REFERENCE: PF-0614 US  
CURRENT APPLICATION NUMBER: US/09/173,581A  
CURRENT FILING DATE: 1998-10-15  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PERL Program  
SEQ ID NO 10  
LENGTH: 1427  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE: -

OTHER INFORMATION: 119819  
US-09-173-581-10

Query Match  
Best Local Similarity 53.7%; Score 35.2; DB 3; Length 1427;  
Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 322 AATGACAGAAATTTTCATTTTCATGGTGACTTCAATGCTGCTGACGCTACGTCCTCAAG 381  
DB 857 AATGACAGATTAAGATCAAGATCAAGATGCGACGCTGCGCAACGCTGCTGCTGCGCAAG 916  
QY 382 AAGGCTTGGAAGACATCCGCTGAGAGCGAGCCCAAGTTCTTTGGCTGATCGGGAC 441  
DB 917 CACTTACGGAACATCCAGACTCGGACGAGTACCGGCGCTCGAGTGCTGATCGCGCC 976  
QY 442 CAAGAGACACCAACGG 457  
DB 977 GAATACGCGCCCGCG 992

## RESULT 14

US-09-420-915-10  
Sequence 10, Application US/09420915  
Patent No. 6264947  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Tang, Y. Tom  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Yue, Henry  
APPLICANT: Guebler, Karl J.  
APPLICANT: Corley, Neil C.  
APPLICANT: Gorgone, Gina  
APPLICANT: Azimzai, Yalda  
APPLICANT: Lu, Aina  
TITLE OF INVENTION: Protein kinase Homologs  
FILE REFERENCE: PF-0614 US  
CURRENT APPLICATION NUMBER: US/09/420,915  
CURRENT FILING DATE: 1999-10-20  
EARLIER APPLICATION NUMBER: US 09/173,581  
EARLIER FILING DATE: 1998-10-15  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PERL Program  
SEQ ID NO 10  
LENGTH: 1427  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE: -  
OTHER INFORMATION: 119819  
US-09-420-915-10

Query Match  
Best Local Similarity 53.7%; Score 35.2; DB 4; Length 1427;  
Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 322 AATGACAGAAATTTTCATTTTCATGGTGACTTCAATGCTGCTGACGCTACGTCCTCAAG 381  
DB 857 AATGACAGATTAAGATCAAGATCAAGATGCGACGCTGCGCAACGCTGCTGCTGCGCAAG 916  
QY 382 AAGGCTTGGAAGACATCCGCTGAGAGCGAGCCCAAGTTCTTTGGCTGATCGGGAC 441  
DB 917 CACTTACGGAACATCCAGACTCGGACGAGTACCGGCGCTCGAGTGCTGATCGCGCC 976  
QY 442 CAAGAGACACCAACGG 457  
DB 977 GAATACGCGCCCGCG 992

## RESULT 15

US-08-924-440-1  
Sequence 1, Application US/08924440  
Patent No. 5871550  
GENERAL INFORMATION:  
APPLICANT: Fitts et al.

TITLE OF INVENTION: MUTANT THERMONOSPORA SPP. CELLULASE  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genencor International, Inc.  
STREET: 925 Page Mill Road  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/924,440

FILING DATE: August 27, 1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Christopher L. Stone

REGISTRATION NUMBER: 35,696

REFERENCE/DOCKET NUMBER: GC388

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 846-7555

TELEFAX: (650) 845-6504

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1293 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-924-440-1

Query Match  
Best Local Similarity 58.0%; Score 32.8; DB 2; Length 1293;  
Matches 58; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 351 CTTCATGCTGCTGAGCTAGCTCCCAAGAGCGCTGGAAGAGATCCGCTGAGAC 410  
DB 1080 CTACACCGGTGACGGCGCCACGACTTCCAGATGGCGGACCGCTACATGATGATGAC 1139  
QY 411 GGACCCCAAGTTGTTGGCTGATCGGGGACCAAGAGAC 450  
DB 1140 GGACCGAAGATCGGTGACCAAGTGAATTAATCTGAC 1179

Search completed: December 9, 2002, 09:19:41  
Job time : 52.0918 secs

**This Page Blank (uspto)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 09:00:58 ; Search time 117.65 seconds  
(without alignments)  
1965.092 Million cell updates/sec

Title: US-09-905-114-1

Perfect score: 592  
Sequence: 1 gagaagctaacggaattc.....gtctgacgaagccctcg 592

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 350425 seqs, 19496369 residues

Total number of hits satisfying chosen parameters: 700850

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA.\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	592	100.0	592	10	US-09-905-114-1
2	592	100.0	671	10	US-09-905-114-3
3	476.8	80.5	1023	10	US-09-962-436-336
4	476.8	80.5	1023	10	US-09-954-456-271
5	476.8	80.5	1023	10	US-09-880-107-3392
6	476.8	80.5	1079	12	US-10-133-065-1
7	476.8	80.5	1079	12	US-10-074-509-1
8	435.2	73.5	1124	12	US-10-133-065-11
9	435.2	73.5	1124	12	US-10-074-509-11
c 10	171.4	25.0	392	10	US-09-880-107-610
11	149.4	25.2	1039	10	US-09-771-078-1
12	149.4	25.2	1039	10	US-09-825-012-2
13	147.8	25.0	783	10	US-09-825-012-4
14	147.8	25.0	858	10	US-09-825-012-6
15	147.8	25.0	1548	10	US-09-825-012-62
16	147.8	25.0	1548	10	US-09-825-012-63
17	147.8	25.0	1548	10	US-09-825-012-65
18	147.8	25.0	1554	10	US-09-825-012-34
19	147.8	25.0	1554	10	US-09-825-012-35

20	147.8	25.0	1554	10	US-09-825-012-37	Sequence 37, Appl
21	147.8	25.0	1557	10	US-09-825-012-64	Sequence 64, Appl
22	147.8	25.0	1560	10	US-09-825-012-72	Sequence 72, Appl
23	147.8	25.0	1560	10	US-09-825-012-73	Sequence 73, Appl
24	147.8	25.0	1560	10	US-09-825-012-75	Sequence 75, Appl
25	147.8	25.0	1560	10	US-09-825-012-77	Sequence 77, Appl
26	147.8	25.0	1560	10	US-09-825-012-78	Sequence 78, Appl
27	147.8	25.0	1560	10	US-09-825-012-79	Sequence 79, Appl
28	147.8	25.0	1563	10	US-09-825-012-36	Sequence 36, Appl
29	147.8	25.0	1566	10	US-09-825-012-67	Sequence 67, Appl
30	147.8	25.0	1566	10	US-09-825-012-68	Sequence 68, Appl
31	147.8	25.0	1566	10	US-09-825-012-70	Sequence 70, Appl
32	147.8	25.0	1569	10	US-09-825-012-74	Sequence 74, Appl
33	147.8	25.0	1575	10	US-09-825-012-69	Sequence 69, Appl
34	147.8	25.0	1578	10	US-09-825-012-81	Sequence 81, Appl
35	147.8	25.0	1578	10	US-09-825-012-82	Sequence 82, Appl
36	147.8	25.0	1578	10	US-09-825-012-84	Sequence 84, Appl
37	147.8	25.0	1584	10	US-09-825-012-39	Sequence 39, Appl
38	147.8	25.0	1584	10	US-09-825-012-40	Sequence 40, Appl
39	147.8	25.0	1584	10	US-09-825-012-42	Sequence 42, Appl
40	147.8	25.0	1587	10	US-09-825-012-83	Sequence 83, Appl
41	147.8	25.0	1590	10	US-09-825-012-91	Sequence 91, Appl
42	147.8	25.0	1590	10	US-09-825-012-92	Sequence 92, Appl
43	147.8	25.0	1590	10	US-09-825-012-94	Sequence 94, Appl
44	147.8	25.0	1593	10	US-09-825-012-41	Sequence 41, Appl
45	147.8	25.0	1596	10	US-09-825-012-86	Sequence 86, Appl

#### ALIGNMENTS

RESULT 1  
US-09-905-114-1  
Sequence 1, Application US/09905114  
Patent No. US2002046745A1  
GENERAL INFORMATION:  
APPLICANT: ZHANG, HUANNIN  
APPLICANT: AX, ROY L  
APPLICANT: BELIN, MARY E  
TITLE OF INVENTION: ISOLATED POLYNUCLEOTIDE SEQUENCES ENCODING A FERTILITY ASSOCIATED  
FILE REFERENCE: 2107070US20  
CURRENT APPLICATION NUMBER: US/09/905,114  
CURRENT FILING DATE: 2001-07-14  
PRIOR APPLICATION NUMBER: US 60/218,140  
PRIOR FILING DATE: 2000-07-14  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 592  
TYPE: DNA  
ORGANISM: Bos sp.  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(591)  
OTHER INFORMATION:  
US-09-905-114-1

Query Match 100.0%; Score 592; DB 10; Length 592;  
Best Local Similarity 100.0%; Pred. No. 1.2e-176;  
Matches 592; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GAGAAGCTAAAGCGAATTCAAGAAAGCATACATACATGATGTTACTCTCGC	60
DB	1	GAGAAGCTAAAGCGAATTCAAGAAAGCATACATACATGATGTTACTCTCGC	60
QY	61	CTTGAAGAAACATATTAAGAACAGTATGCTTCTCTATTAAGAAAGCTAGTGTCT	120
DB	61	CTTGAAGAAACATATTAAGAACAGTATGCTTCTCTATTAAGAAAGCTAGTGTCT	120
QY	121	GTTAAACAAAGCTACTCTTACACGACTATCAAGCTGAGAGAGAGAGATGTTTCCAGG	180
DB	121	GTTAAACAAAGCTACTCTTACACGACTATCAAGCTGAGAGAGAGATGTTTCCAGG	180





Oy 361 GCGTCAGCTACGTCCTCCCAAGAGGCGCTGGAAGACATCCGCTGAGACGAGCCCAAG 420  
 Db 601 GCGTCAGCTACGTCCTCCCAAGAGGCGCTGGAAGACATCCGCTGAGACGAGCCCAAG 660  
 Oy 421 TTCGTTGGCTGATCGGGACCAAGAGACACCAAGCTCAAGAAAGACAACTGCGCC 480  
 Db 661 TTTGTTGGCTGATCGGGACCAAGAGACACCAAGCTCAAGAAAGACAACTGCGCA 720  
 Oy 481 TATGACAGGATGCTGCTTAGAGGACAAATATGTCACATCTGCTGCTCTCAATCAAC 540  
 Db 721 TATGACAGGATGCTGCTTAGAGGACAAATATGTCACATCTGCTGCTCTCAATCAAC 780  
 Oy 541 CTCGCTTTGATTTCCAGAAAGCTTACAGGTTGCTGAATCGAAGGCGCTGG 592  
 Db 781 AGTGTTTGGACTTCCAGAAAGCTTACAGGTTGCTGAATCGAAGGCGCTGG 832

RESULT 4

US-09-954-456-271  
 : Sequence 271, Application US/09954456  
 : Patent No. US20020115057A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Young, Paul  
 : TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc  
 : FILE REFERENCE: 689290-76  
 : CURRENT APPLICATION NUMBER: US/09/954,456  
 : CURRENT FILING DATE: 2001-09-18  
 : PRIOR APPLICATION NUMBER: US/60/233,617  
 : PRIOR FILING DATE: 2000-09-18  
 : PRIOR APPLICATION NUMBER: US/60/234,052  
 : PRIOR FILING DATE: 2000-09-20  
 : PRIOR APPLICATION NUMBER: US/60/234,923  
 : PRIOR FILING DATE: 2000-09-25  
 : PRIOR APPLICATION NUMBER: US/60/235,134  
 : PRIOR FILING DATE: 2000-09-25  
 : PRIOR APPLICATION NUMBER: US/60/235,637  
 : PRIOR FILING DATE: 2000-09-26  
 : PRIOR APPLICATION NUMBER: US/60/235,638  
 : PRIOR FILING DATE: 2000-09-26  
 : PRIOR APPLICATION NUMBER: US/60/235,711  
 : PRIOR FILING DATE: 2000-09-27  
 : PRIOR APPLICATION NUMBER: US/60/235,720  
 : PRIOR FILING DATE: 2000-09-27  
 : PRIOR APPLICATION NUMBER: US/60/235,840  
 : PRIOR FILING DATE: 2000-09-27  
 : PRIOR APPLICATION NUMBER: US/60/235,863  
 : PRIOR FILING DATE: 2000-09-27  
 : NUMBER OF SEQ ID NOS: 2276  
 : SOFTWARE: PatentIn version 3.0  
 : SEQ ID NO 271  
 : LENGTH: 1023  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens  
 : US-09-954-456-271

Query Match 80.5%; Score 476.8; DB 10; Length 1023;  
 Best Local Similarity 87.8%; Pred. No. 3.1e-140;  
 Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

Oy 1 GAGAAAGCTAAAGCGAAATTCAGAAAGGATTAACATACATCATGATGATGATGCTCGC 60  
 Db 241 GAGAAAGCTAAAGCGAAATTCAGAAAGGATTAACATACATCATGATGATGATGCTCGC 300  
 Oy 61 CTGGAAGAAACACATATTAAGAAAGATGCTTTCTCTATTAAGAAAGGATGCTCT 120  
 Db 301 CTGGAAGAAACACATATTAAGAAAGATGCTTTCTCTATTAAGAAAGGATGCTCT 360  
 Oy 121 GTAAGAAAGGATGCTCTCTACCAAGCATATCAGGCTGGAGAGCGCAATGTTTCCAG 180  
 Db 361 GTAAGAAAGGATGCTCTCTACCAAGCATATCAGGATGAGAGCGCAATGTTTCCAG 420

Oy 181 GAAACCTTTGTGCTGCTTCCAGTCACCTTACACCGCTGTCAAGGACCTTGTGATGTC 240  
 Db 421 GAGCCCTTTGTGCTGCTTCCAGTCACCTTACACCGCTGTGTCAAGGACCTTGTGATGTC 480  
 Oy 241 CCCCTGCAACCAACCTTACAGATCTCCGTTAGAGAAATGATGATGCTGCTGATGTC 300  
 Db 481 CCCCTGCAACCAACCTTACAGATCTCCGTTAGAGAAATGATGATGCTGCTGATGTC 540  
 Oy 301 ACAGATGTAAACGTCGCTGGAATGAGAAATTTTATTTTCAATGAGGCTCAATGCT 360  
 Db 541 ACAGATGTAAACGTCGCTGGAATGAGAAATTTTATTTTCAATGAGGCTCAATGCT 600  
 Oy 361 GCGTCAGCTACGTCCTCCCAAGAGGCGCTGGAAGACATCCGCTGAGACGAGCCCAAG 420  
 Db 601 GCGTCAGCTACGTCCTCCCAAGAGGCGCTGGAAGACATCCGCTGAGACGAGCCCAAG 660  
 Oy 421 TTCGTTGGCTGATCGGGACCAAGAGACACCAAGCTCAAGAAAGACAACTGCGCC 480  
 Db 661 TTTGTTGGCTGATCGGGACCAAGAGACACCAAGCTCAAGAAAGACAACTGCGCA 720  
 Oy 481 TATGACAGGATGCTGCTTAGAGGACAAATATGTCACATCTGCTGCTCTCAATCAAC 540  
 Db 721 TATGACAGGATGCTGCTTAGAGGACAAATATGTCACATCTGCTGCTCTCAATCAAC 780  
 Oy 541 CTCGCTTTGATTTCCAGAAAGCTTACAGGTTGCTGAATCGAAGGCGCTGG 592  
 Db 781 AGTGTTTGGACTTCCAGAAAGCTTACAGGTTGCTGAATCGAAGGCGCTGG 832

RESULT 5

US-09-880-107-3392  
 : Sequence 3392, Application US/09880107  
 : Patent No. US20020142981A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Horne, Darci T.  
 : APPLICANT: Vockley, Joseph G.  
 : APPLICANT: Scherf, Uwe  
 : APPLICANT: Gene Logic, Inc.  
 : TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
 : FILE REFERENCE: 44921-5028-NO  
 : CURRENT APPLICATION NUMBER: US/09/880,107  
 : CURRENT FILING DATE: 2001-06-14  
 : PRIOR APPLICATION NUMBER: US 60/211,379  
 : PRIOR FILING DATE: 2000-06-14  
 : PRIOR APPLICATION NUMBER: US 60/237,054  
 : PRIOR FILING DATE: 2000-10-02  
 : NUMBER OF SEQ ID NOS: 3950  
 : SOFTWARE: PatentIn Ver. 2.1  
 : SEQ ID NO 3392  
 : LENGTH: 1023  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens  
 : FEATURE:  
 : OTHER INFORMATION: Genbank Accession No. US20020142981A1 US6814  
 : US-09-880-107-3392

Query Match 80.5%; Score 476.8; DB 10; Length 1023;  
 Best Local Similarity 87.8%; Pred. No. 3.1e-140;  
 Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

Oy 1 GAGAAAGCTAAAGCGAAATTCAGAAAGGATTAACATACATCATGATGATGATGCTCGC 60  
 Db 241 GAGAAAGCTAAAGCGAAATTCAGAAAGGATTAACATACATCATGATGATGATGCTCGC 300  
 Oy 61 CTGGAAGAAACACATATTAAGAAAGATGCTTTCTCTATTAAGAAAGGATGCTCT 120  
 Db 301 CTGGAAGAAACACATATTAAGAAAGATGCTTTCTCTATTAAGAAAGGATGCTCT 360  
 Oy 121 GTAAGAAAGGATGCTCTCTACCAAGCATATCAGGCTGGAGAGCGCAATGTTTCCAG 180  
 Db 361 GTAAGAAAGGATGCTCTCTACCAAGCATATCAGGATGAGAGCGCAATGTTTCCAG 420  
 Oy 181 GAAACCTTTGTGCTGCTTCCAGTCACCTTACACCGCTGTCAAGGACCTTGTGATGTC 240

Db 421 GAGCCCTTTGGTCTGGTCAATCTCCACACATGCTGTCAGAGACTTCGTGATATC 480  
Qy 241 CCCGTGCACACACCCCTGAGACATCCGTTAGAGAGATGATGAGCTGGCTGATGCTAC 300  
Db 481 CCCCTGCACACACCCACAGACATCCGTTAAGAGATCGATGAGTTGGTTGAGGCTAC 540  
Qy 301 ACAGATGTGAACGTCGCTGGAATCGAGAGATTTCAATTTTCATGGGTGACTTCATGCT 360  
Db 541 ACGGACGTGAACACACCGTGGAGGCGGAGAAATTTCAATTTTCATGGGTGACTTCATGCC 600  
Qy 361 GGCTGCAGCTACGTCCTCCCAAGAGCCCTGGAAGGACATCCGCTGAGGACGAGACCCCAAG 420  
Db 601 GGCTGCAGCTACGTCCTCCCAAGAGCCCTGGAAGACATCCGCTGAGGACGAGACCCCAAG 660  
Qy 421 TTCGTTGGCTGATCGGGGACCAAGAGACACCGTCAAGAGAGACAAAACCTGGGCC 480  
Db 661 TTTGTTGGCTGATCGGGGACCAAGAGACACCGTCAAGAGAGACAAAACCTGTGCA 720  
Qy 481 TATGACAGATCGCTGTTAGAGGACAAATATTGTCAACTCTGGTGGTCTCAATCAAC 540  
Db 721 TATGACAGATGCTGTTAGAGGACAAAGAAATCGTCAGTTCTGTTGTTCCCAAGTCAAAC 780  
Qy 541 CTCGCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTCG 592  
Db 781 AGTGTGTTGACTCCAGAAAGCTTACAGCTGACTGAGAGAGGCCCTCG 832

RESULT 6  
US-10-133-065-1  
; Sequence 1, Application US/10133065  
; Patent No. US20020123122A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; Baron, Will F.  
; TITLE OF INVENTION: HUMAN DNASE  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; City: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/133,065  
; FILING DATE: 26-Apr-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/643,520  
; FILING DATE: 22/08/2000  
; APPLICATION NUMBER: 08/794827  
; FILING DATE: 04-Feb-1997  
; APPLICATION NUMBER: 60/109796  
; FILING DATE: 05-Feb-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Evans, David W  
; REGISTRATION NUMBER: \*See attached Limited Recognition under 37  
; REFERENCE/DOCKET NUMBER: C.F.R. 10.9(b)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-1739  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1079 base pairs  
; TYPE: Nucleic Acid  
; STRANDEDNESS: Single

; TOPOLOGY: Linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-133-065-1  
Query Match 80.5%; Score 476.8; DB 12; Length 1079;  
Best Local Similarity 87.8%; Pred. No. 3.2e-140;  
Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;  
Qy 1 GAGAAGCTAAACGGAATTCAGAGAAAGGCATACATACAACTATGTGATTAGCTCTCGC 60  
Db 287 GAGAAGCTGAACGAAATTCAGAGAGAGGCATACGTTACAACTATGTGATTAGCTCTCG 346  
Qy 61 CTTGGAAGAAACATATATAAGACACATATGCTTTCTTATAAGAAAGCTAGTGTCT 120  
Db 347 CTTGGAAGAAACATATATAAGAAACATATGCTTTCTTATAAGAAAGCTAGTGTCT 406  
Qy 121 GTAAACAAAGAGTACCTCTACACGACATATCAGGCTGGAGACGAGATGTGTTTCCAGG 180  
Db 407 GTGAGAGAGAGTATATCACTACCATGATATCAGGATGGAGACGAGATGTGTTTCCAGG 466  
Qy 181 GAACCCCTTTGGTCTGGTTCAGTCCAGTACACCCCTGACACCCCTGTCAGAGACTTCGTGATTGC 240  
Db 467 GAGCCCTTTGGTCTGGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 526  
Qy 241 CCCCTGCACACACCCCTGAGACATCCGTTAGAGAGATTCAGTGGCTGATGCTTAC 300  
Db 527 CCCCTGCACACACCCCTGAGACATCCGTTAAGGAGATCGATGAGTGGTGGATGCTTAC 586  
Qy 301 ACAGATGTGAACCGTCTCGTGGATGAGAGAAATTTCAATTTTCATGGGTGACTTCATGCT 360  
Db 587 ACGGACGTGAACACACCGCTGGAAGGCGGAGAAATTTCAATTTTCATGGGTGACTTCATGCT 646  
Qy 361 GGCTGCAGCTACGTCCTCCCAAGAGGCTCGAAGGACATCCGCTGAGAGACGAGACCCCAAG 420  
Db 647 GGCTGCAGCTACGTCCTCCCAAGAGGCTCGAAGGACATCCGCTTGGAGACTGACCCCAAG 706  
Qy 421 TTGCTTTGGCTGATCGGGGACCAAGAGACACCGTCAAGAGAGACAAAACCTGCTCAATCAAC 480  
Db 707 TTTGTTGGCTGATCGGGGACCAAGAGAGACACCGTCAAGAGAGACAAAACCTGCTCA 766  
Qy 481 TATGACAGATCGTGTCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCTCTCAATCAAC 540  
Db 767 TATGACAGATGTTGCTTACAGGACAAATCGTCAGTTCTGTTGTTCCCAAGTCAAAC 826  
Qy 541 CTGCTCTTTGATTTCCAGAAAGCTTACAGGTTGTCTGAAATCGAAGGCCCTCG 592  
Db 827 AGTGTGTTTGAATTTCCAGAAAGCTTACAGCTGACTGAGAGAGGCCCTCG 878

RESULT 7  
US-10-074-509-1  
; Sequence 1, Application US/10074509  
; Patent No. US20020142437A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; Baron, Will F.  
; TITLE OF INVENTION: HUMAN DNASE  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; City: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/074,509  
; FILING DATE: 11-Feb-2002

CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/643,520  
FILING DATE: 22/08/2000  
APPLICATION NUMBER: 08/794827  
FILING DATE: 04-Feb-1997  
APPLICATION NUMBER: 60/109796  
FILING DATE: 05-Feb-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Johnston, Sean A.  
REGISTRATION NUMBER: 35,910  
REFERENCE/DOCKET NUMBER: P1000R1C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-3562  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1079 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-074-509-1  
Query Match 80.5%; Score 476.8; DB 12; Length 1079;  
Best Local Similarity 87.8%; Pred. No. 3.2e-140;  
Matches 520; Conservative 0; Mismatches 72; Indels 0; Gaps 0;  
QY 1 GAGAGCTAAACGGAATTCAGAAAAGGACATACATACATATGATTAGCTCTGC 60  
DB 287 GAGAACTGAACGAAATTCAGAGAGGCGATACGACAACTATGTAGTACCTCGG 346  
QY 61 CTGGAGAAACATATAAGAACAGTACCTTCTCTATAAAGAAAGTACGTCT 120  
DB 347 CTGGAGAAACATATAAGAACATATGCTTCTCTACAGGAAAGCTGTGTCT 406  
QY 121 GTAAAACAAAGCTACCTCTACAGACTATCAGCTGAGACGAGATGTTTCCAG 180  
DB 407 GTGAAGAGAGTATATCACTACATGACTATCAGATGAGACCCAGATGTGTTCCAG 466  
QY 181 GAACCTTTGTGTGTTGTTCCAGTACCTTACACCGCTGTCAAGACCTTGATTGTC 240  
DB 467 GAGCCCTTTGTGTGTTGTTCCAGTACCTTACACCGCTGTCAAGACCTTGATTGTC 526  
QY 241 CCCCTGACACACCCCTGAGACATCCGTTAGAGATGATGAGCTGCTGATGCTAC 300  
DB 527 CCCCTGACACACCCCTGAGACATCCGTTAGAGATGATGAGCTGCTGATGCTAC 586  
QY 301 ACAGATGTGAACGTCGTGGAATGAGAGATTTTCATTTGAGTGAATCAATGCT 360  
DB 587 ACAGATGTGAACGTCGTGGAATGAGAGATTTTCATTTGAGTGAATCAATGCT 646  
QY 361 GCGTGAAGTACGTCCTCCAGAAAGGCTGTGAGAGACATCCGCTGAGACGACCCCAAG 420  
DB 647 GCGTGAAGTACGTCCTCCAGAAAGGCTGTGAGAGACATCCGCTGAGACGACCCCAAG 706  
QY 421 TTGCTTTGCTGATCGGGGACCAAGAGACACACAGCTGCAAGAAAGCAAAATGCGCC 480  
DB 707 TTGCTTTGCTGATCGGGGACCAAGAGACACACAGCTGCAAGAAAGCAAAATGCGCC 766  
QY 481 TATGACAGATCGTGTGAGAGACAAATATTTGCACTCTGCTGCTCTCATCAAA 540  
DB 767 TATGACAGATCGTGTGAGAGACAAATATTTGCACTCTGCTGCTCTCATCAAA 826  
QY 541 CTCGCTTTGATTTCCAGAAAGCTTACAGGTTGCTGAATGGAAGGCTCTGG 592  
DB 827 AGTGTTTTCAGCTTCCAGAAAGCTTACAGGTTGCTGAATGGAAGGCTCTGG 878

RESULT 9  
US-10-133-065-11  
Sequence 11, Application US/10133065  
Patent No. US20020123122A1

GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
TITLE OF INVENTION: HUMAN DNASE  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/133,065  
FILING DATE: 26-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/643,520  
FILING DATE: 22/08/2000  
APPLICATION NUMBER: 08/794827  
FILING DATE: 04-Feb-1997  
APPLICATION NUMBER: 60/109796  
FILING DATE: 05-Feb-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Evans, David W  
REGISTRATION NUMBER: \*See attached Limited Recognition under 37 C.F.R. 10.9(b)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-1739  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1124 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 11:  
US-10-133-065-11  
Query Match 73.5%; Score 435.2; DB 12; Length 1124;  
Best Local Similarity 83.4%; Pred. No. 4.2e-127;  
Matches 494; Conservative 0; Mismatches 98; Indels 0; Gaps 0;  
QY 1 GAGAGCTAAACGGAATTCAGAAAAGGACATACATACATATGATTAGCTCTGC 60  
DB 404 GAGAGCTGAATGAAATTCAGAAAAGGACACAACTATGATGATTAGCTCTGCA 463  
QY 61 CTGGAGAAACATATAAGAACAGTATGCTTCTCTATAAAGAAAGCTAGTGTCT 120  
DB 464 CTGGAGAAACATATAAGAACAGTATGCTTCTCTACAGAGAGAGCTGTGTCT 523  
QY 121 GTAAAACAAAGCTACCTCTACAGACTATCAGCTGAGACCCAGATGTTTCCAG 180  
DB 524 GTGAAGAGAGTATATCACTACATGACTATCAGATGAGAGACACACAGCTTTCCAG 583  
QY 181 GAACCTTTGTGTGTTGTTCCAGTACCTTACACCGCTGTCAAGACCTTGATTGTC 240  
DB 584 GAGCCCTTTGTGTGTTGTTCCAGTACCTTACACCGCTGTGTCAAGACCTTGATTGTC 643  
QY 241 CCCCTGACACACCCCTGAGACATCCGTTAGAGATGATGAGCTGCTGATGCTAC 300  
DB 644 CCCCTGACACACCTCCGAGACCTCGTTAAAGATATGATGCTGTGCTATGCTAC 703  
QY 301 ACAGATGTGAACGTCGTGGAATGAGAGATTTTCATTTGAGTGAATCAATGCT 360  
DB 704 ACAGATGTGAACGTCGTGGAATGAGAGATTTTCATTTGAGTGAATCAATGCT 763

RESULT 9  
US-10-074-509-11  
Sequence 11, Application US/10074509  
Patent No. US20020142437A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
Baron, Will F.  
TITLE OF INVENTION: HUMAN DNASE  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/074,509  
FILING DATE: 11-Feb-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/643,520  
FILING DATE: 22/08/2000  
APPLICATION NUMBER: 08/794827  
FILING DATE: 04-Feb-1997  
APPLICATION NUMBER: 60/109796  
FILING DATE: 05-Feb-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Johnston, Sean A.  
REGISTRATION NUMBER: 35,910  
REFERENCE/DOCKET NUMBER: PI000R1C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-3562  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1124 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 11:  
US-10-074-509-11

Qy	61	CTTTGGAGAAACACATATAAAGACAGTATGCTTTCTTCTATAAAGAAAGCTAGTGTCT	120
Db	464	CTTTGGAGAAACACGTCTAAAGAGCAGTATGCTTCGCTCTACAGAGNAGCTGTGTCT	523
Qy	121	GTAAACAAGCTTACCTCTACCAACGACTATCAGGCTGGAGACGCAGATGTGTTTTCCAGG	180
Db	524	GTGAAGACAAAATACCACCTACCATGACTATCAGGATGGAGACACAGACGTGTTTTCCAGG	583
Qy	181	GAACCCCTTTGTGTCTGTGTTCCAGTTCACCTACACGCTGTCAAGGACTTTCGTGATGTGC	240
Db	584	GAGCCCTTTGTGTGTTGTGTTCCAACTCCCCCTTTACTGCTGTCAAGGACTTCGTGATGTGC	643
Qy	241	CCCTGCACACCAACCCTCGTAGACATCCGTTTAGAGAGATTGATCAGCTGCCTGATGTCTAC	300
Db	644	CCCTTGACACAACTCCGAGACCTCCGTTAAAGAGATAGATGAGCTGTCATGCTCTAC	703
Qy	301	ACAGATGTGAAACGTGCTGTGGAAATTCAGAGAAATTTCTATTTTCATGGTGACTTCAATGCT	360
Db	704	ACGGATGTGAAAGCCAGTGGAAAGACAGAGAAATTTCTATCTTCATGGGTGATTTTCAACGGC	763
Qy	361	GGCTGCAGCTACGTGCCCAAGAAAGGCTCGAAGGACATCCGCTGAGGACGACCCCAAG	420
Db	764	GGCTGTAGCTATGTGCCCAAGAGGCTCGCAGAACATTCGTTTGAGGACGACCCCAAG	823
Qy	421	TTCGTTTGGCTGATCGGGGACCAAGAGGACACACCGTCAAGAAGAGCAAAACTGCGGCC	480
Db	824	TTTGTGTTGGCTGATTGGGACCAAGAGGACACTACCGTCAAGAAGAGTACCAGCTGTGCGC	883
Qy	481	TATGACAGGATCGTCTTAGAGGACAAAATATTGTCAACTCTGGTGGTCCCTCAATCAAAAC	540
Db	884	TATGACAGGATTGTGCTTTTGTGACAGAGATAGTCAATCTCGTGGTTCCCGTCCAGT	943
Qy	541	CTCGCTTTGATTTCCAGAAAGCTTACAGGTGTCTGAAATCGAAGGCCCTGG	592
Db	944	GCCTCTTTCAGCTTTCAGAAGCTTATGACTTCTCTGAGGAGAGGCCCTGG	995

```

RESULT 10
US-09-880-107-630/c
; Sequence 630, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 630
; LENGTH: 392
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA279676
US-09-880-107-630

```

	Query Match	29.0%	Score 171.4;	DB 10;	Length 392;
	Best Local Similarity	86.0%	Pred. No. 2.8e-44;		
	Matches 190;	Conservative	0;	Mismatches 31;	Indels 0;
	Gaps				
QY	372	CGTCCCAAGAGCGCTGGAAAGACATCCGCTGTAGAGACGACCCCAAGTTGCTTTGGCT	431		
Db	392	CGTCCCAAGAGCGCTGGAAAGAACATCCGCTTGTAGAGACTGACCCCAAGTTTGTTTGGCT	333		
QY	432	GATCGGGACCAAGAGGACACCCAGGTCAAGAGAGACCAAACTGTGCGCTTTGACAGGAT	491		

Db 332 GATCGGGGACCAAGAGACACCGGTGAAGAGAGACCAATGTGCATATGACAGGAT 273  
Qy 492 CGTCTTAGAGGACAAATATTTCAACTCTGGTGTCTCTCAATCAAACTCTCTTTTGA 551  
Db 272 TGTGCTTAGAGGACAAATATCTGCTGTTGTTCCCAAGTCAACAGTGTTTTGA 213  
Qy 552 TTTCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCCCTGG 592  
Db 212 CTTCCAGAAAGCTTACAAGCTGACTGAAAAGGAGGCCCTGG 172

## RESULT 11

US-09-771-078-1  
; Sequence 1, Application US/09771078  
; Patent No. US20020034727A1  
; GENERAL INFORMATION:  
; APPLICANT: Mtsny, Randall J.,  
; APPLICANT: Daugherty, Ann L.,  
; APPLICANT: Patapoff, Thomas W.  
; TITLE OF INVENTION: DNase Compaction Assay  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/771,078  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/971,019  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Glaister, Debra J.  
; REGISTRATION NUMBER: 33,988  
; REFERENCE/DOCKET NUMBER: 792  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-3177  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1039 bases  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

## US-09-771-078-1

Query Match 25.2%; Score 149.4; DB 10; Length 1039;  
Best Local Similarity 56.7%; Pred. No. 4.1e-37;  
Matches 317; Conservative 0; Mismatches 236; Indels 6; Gaps 2;  
Qy 34 ACATACAACTATGTGATTAGCTCTCGCCCTTGGAAAGAAACACATATAAAGAACAGTATGCC 93  
Db 409 ACCTATCACTACGTGTGCTAGTGAGCCACTGGGACGAAACAGCTATTAAGGAGCGCTACCTG 468  
Qy 94 TTTCTCTATAAAGAAAGCTAGTGTCTGTAAACAAAGCTACTCTTACCAGCA---CTAT 150  
Db 469 TTCGTGTACAGGCTGACCAAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGCTGC 528  
Qy 151 CAGGCTGGAGACGACATGTGTTTCCAGGGAACCTTTGTGCTGTGTTCCAGTCACCC 210  
Db 529 GAGCCCTGGGGAAACGACACCTTCAACCGAGAGCCAGCCATTTGTCCAGGTTCTTCTCCCGG 588  
Qy 211 TACACCGCTGTCAAGGACTTCGTGATTGTTCCCTGCAACACCCCTCGAGCATCCGTT 270  
Db 589 TTCACAGAGGTCAGGAGTTTGCCATTGTTCCCTGCACTGCGGCCCGCGGGGACGAGTA 648  
Qy 271 AGAGAGATTGATGAGCTGGCTGATGTCTACAGATGTGAACGCTGCTGGAATGCAGAG 330

## US-09-771-078-1

Query Match 25.2%; Score 149.4; DB 10; Length 1039;  
Best Local Similarity 56.7%; Pred. No. 4.1e-37;  
Matches 317; Conservative 0; Mismatches 236; Indels 6; Gaps 2;  
Qy 34 ACATACAACTATGTGATTAGCTCTCGCCCTTGGAAAGAAACACATATAAAGAACAGTATGCC 93  
Db 409 ACCTATCACTACGTGTGCTAGTGAGCCACTGGGACGAAACAGCTATTAAGGAGCGCTACCTG 468  
Qy 94 TTTCTCTATAAAGAAAGCTAGTGTCTGTAAACAAAGCTACTCTTACCAGCA---CTAT 150  
Db 469 TTCGTGTACAGGCTGACCAAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGCTGC 528  
Qy 151 CAGGCTGGAGACGACATGTGTTTCCAGGGAACCTTTGTGCTGTGTTCCAGTCACCC 210  
Db 529 GAGCCCTGGGGAAACGACACCTTCAACCGAGAGCCAGCCATTTGTCCAGGTTCTTCTCCCGG 588

Qy 211 TACACCGCTGTCAAGGACTTCGTGATTGTTCCCTGCAACACCCCTGAGACATCCGTT 270  
Db 589 TTCACAGAGGTCAGGAGTTTGCCATTGTTCCCTGCACTGCGGCCCGGGGACGAGTA 648  
Qy 271 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAACGCTCGCTGGAATGCAGAG 330  
Db 649 GCGGAGATCGACGCTCTCTATGACGCTACTCTGATGTCCAAGAGAAATGGGCGCTGGAG 708  
Qy 331 AATTTTCATTTTCATGGGTGACTTCAATGCTGGCTGAGCTACGTCCTCCCAAGAGGCTGG 390  
Db 709 GAGCTCATGTTGATGGCGACTTCAATGCGGCTGAGCTATGTGAGACCTCCAGTGG 768  
Qy 391 AAGGACATCGGCTGAGGACGACCCCAAGTTGCTTGGCTGATCGGGGACCAAGAGGAC 450  
Db 769 TCATCCATCCGCTGTGGACAAAGCCCCACCTTCCAGTGGCTGATCCCCGACAGCGCTGAC 828  
Qy 451 ACCAGGTCAAGAAGAGACAAACATCGCGCTATGACAGGATCGTCTTATAGAGACAAAAT 510  
Db 829 ACCACAG---CTACCCACGACACTGTGCTATGACAGGATCGTGTTCAGGGATGCTG 885  
Qy 511 ATTGTCAACTCTGGTGTCTCTCAATCAAACTCGCTCTTTGATTTCCAGAAAGCTTACAGG 570  
Db 886 CTCGAGGCGCGTGTGTTCCCGACTCGGCTCTTCCCTTTAACTTCCAGGCTGCTATGGC 945  
Qy 571 TTGCTGTAATCGAAGGCC 589  
Db 946 CTGAGTGACCAACTGGCCC 964

## RESULT 12

US-09-825-012-2  
; Sequence 2, Application US/09825012  
; Patent No. US20020122798A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Robert  
; TITLE OF INVENTION: Compounds for Targeting  
; FILE REFERENCE: 43191-256808  
; CURRENT APPLICATION NUMBER: US/09/825,012  
; CURRENT FILING DATE: 2001-04-03  
; PRIOR APPLICATION NUMBER: US 60/237,159  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: GB 0008049.9  
; PRIOR FILING DATE: 2000-04-03  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 1039  
; TYPE: DNA  
; ORGANISM: Homo sapiens

## US-09-825-012-2

Query Match 25.2%; Score 149.4; DB 10; Length 1039;  
Best Local Similarity 56.7%; Pred. No. 4.1e-37;  
Matches 317; Conservative 0; Mismatches 236; Indels 6; Gaps 2;

Qy 34 ACATACAACTATGTGATTAGCTCTCGCCCTTGGAAAGAAACACATATAAAGAACAGTATGCC 93  
Db 409 ACCTATCACTACGTGTGCTAGTGAGCCACTGGGACGAAACAGCTATTAAGGAGCGCTACCTG 468  
Qy 94 TTTCTCTATAAAGAAAGCTAGTGTCTGTAAACAAAGCTACTCTTACCAGCA---CTAT 150  
Db 469 TTCGTGTACAGGCTGACCAAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGCTGC 528  
Qy 151 CAGGCTGGAGACGACATGTGTTTCCAGGGAACCTTTGTGCTGTGTTCCAGTCACCC 210  
Db 529 GAGCCCTGGGGAAACGACACCTTCAACCGAGAGCCAGCCATTTGTCCAGGTTCTTCTCCCGG 588  
Qy 211 TACACCGCTGTCAAGGACTTCGTGATTGTTCCCTGCAACACCCCTCGAGCATCCGTT 270  
Db 589 TTCACAGAGGTCAGGAGTTTGCCATTGTTCCCTGCACTGCGGCCCGCGGGGACGAGTA 648  
Qy 271 AGAGAGATTGATGAGCTGGCTGATGTCTACAGATGTGAACGCTGCTGGAATGCAGAG 330

Db 649 GCCGAGTCAGAGCTCTCTATGACGTCTACCTGATGTCCAGAGAAATGGGGCTTGGAG 708  
Qy 331 AATTTCAATTTTCATGAGTACTTCAATGCTGCTGACGTACCTCCCAAGAGGCTTGG 390  
Db 709 GACGTCATGTTGATGGGCGAATTCATATGCGGCTGACAGTATGTAGACCTTCCCACTGG 768  
Qy 331 AAGGACATCCGCTGAGAGCGGACCCCAAGTTCGTTTGGCTGATGGGAGCCAAAGAGAC 450  
Db 769 TCATCCATCCGCTGTGAGCAAGCCCACTTCCAGTGTCTATCCCAAGCGCTGAC 828  
Qy 451 ACCAGGTCAGAGAGCAAACTGCGCTATGACAGAGTGTGCTTACAGAGCAAAAT 510  
Db 829 ACCAGAG---CTACACCCAGCAGCTGTGCTATGACAGATGTGTTGACAGGATGCTG 885  
Qy 511 ATTGTCAACTGTGTGCTTCAATCAAACTGCTTGTGATTTCCAGAAAGCTTACAG 570  
Db 886 CTCGAGGCGCGCTGTGCTTCCGACCTGCGCTTCCCTTTAACTTCCAGGCTGCTATGAC 945  
Qy 571 TTGTCTGAATGGAAGGCC 589  
Db 946 CTGAGTGAACCAACTGGCCC 964

## RESULT 13

US-09-825-012-4  
; Sequence 4, Application US/09825012  
; Patent No. US2002012798A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Robert  
; TITLE OF INVENTION: Compounds for Targeting  
; FILE REFERENCE: 43191-256808  
; CURRENT FILING DATE: 2001-04-03  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: US 60/237,159  
; PRIOR FILING DATE: 2000-04-03  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 783  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-825-012-4

Query Match 25.0%; Score 147.8; DB 10; Length 783;  
Best Local Similarity 56.5%; Pred. No. 1.1e-36;  
Matches 316; Conservative 0; Mismatches 237; Indels 6; Gaps 2;

Qy 34 ACATCAACTATGTATGATGCTCTGCGCTTGGAGAAACATATTAAGAAAGATATGCC 93  
Db 184 ACCTATCACTACGTGTGATGAGGCACTGGGAGCGAAAGCTATTAAGAGGCTTACCTG 243  
Qy 94 TTTCTCTAATAAGAAAGCTAGTGTCTGTAAACAAGCTACTCTTACACAGA---CTAT 150  
Db 244 TTCGTGTACAGGCTGACAGGTGTCTGCGGTGAGACAGTACTATACATATATGGCTGC 303  
Qy 151 CAGGCTGAGAGCGAGATGTGTTTCCAGGGAACCTTTGGTGTGCTTCAGTACCC 210  
Db 304 GAGCCTCGGGGAAGCACTTCAACGAGAGCCAGCATGTGAGGTTCTTCCCGG 363  
Qy 211 TACACCGCTGTCAAGAGACTTGTGTTCCCTCCCTGCACACCACTCCCTGAGCATCCGT 270  
Db 364 TTCACAGGTCAGGAGATTTGCAATGTTCCCTGCAATGCGGCCGCGGAGCACTA 423  
Qy 271 AGAGAGATTTGATGAGCTGGCTGATCTTCAACAGATGTGAAAGCTGCTGGAATGCAAG 330  
Db 424 GCCGAGATCGAGCTCTCTATGACGTCTACCTGATGTCCAAAGAAATGGGGCTTGGAG 483  
Qy 331 AATTTCAATTTTCATGAGTACTTCAATGCTGCTGCAAGTATGCTCCCAAGAGGCTTGG 390  
Db 484 GACGTCATGTTGATGGGCGAATTCATATGCGGCTGACAGTATGTGAGAACCTCCAGTGG 543

Qy 391 AAGGACATCCGCTGAGAGCGAGCCCAAGTTTCTTTGGCTGATCGGGAGCCAAAGAGAC 450  
Db 544 TCATCCATCCGCTGTGAGCAAGGCCCACTTCCAGTGGCTGATCCCGACAGGCGCTGAC 603  
Qy 451 ACCAGGTCAGAGAGCAAACTGCGCTATGACAGAGTGTGCTTACAGAGCAAAAT 510  
Db 604 ACCAGAG---CTACACCCAGCAGCTGTGCTTATGACAGATGTGTTGACAGGATGCTG 660  
Qy 511 ATTGTCAACTGTGTGCTTCAATCAAACTGCTTGTGATTTCCAGAAAGCTTACAG 570  
Db 661 CTCGAGGCGCGCTGTGCTTCCGACCTGCGCTTCCCTTTAACTTCCAGGCTGCTATGAC 720  
Qy 571 TTGTCTGAATGGAAGGCC 589  
Db 721 CTGAGTGAACCAACTGGCCC 739

## RESULT 14

US-09-825-012-6  
; Sequence 6, Application US/09825012  
; Patent No. US2002012798A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Robert  
; TITLE OF INVENTION: Compounds for Targeting  
; FILE REFERENCE: 43191-256808  
; CURRENT FILING DATE: 2001-04-03  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: US 60/237,159  
; PRIOR FILING DATE: 2000-04-03  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 858  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-825-012-6

Query Match 25.0%; Score 147.8; DB 10; Length 858;  
Best Local Similarity 56.5%; Pred. No. 1.2e-36;  
Matches 316; Conservative 0; Mismatches 237; Indels 6; Gaps 2;

Qy 34 ACATCAACTATGTATGATGCTCTGCGCTTGGAGAAACATATTAAGAAAGATATGCC 93  
Db 259 ACCTATCACTACGTGTGATGAGGCACTGGGAGCGAAAGCTATTAAGAGGCTTACCTG 318  
Qy 94 TTTCTCTAATAAGAAAGCTAGTGTCTGTAAACAAGCTACTCTTACACAGA---CTAT 150  
Db 319 TTCGTGTACAGGCTGACAGGTGTCTGCGGTGAGACAGTACTATACATATATGGCTGC 378  
Qy 151 CAGGCTGAGAGCGAGATGTGTTTCCAGGGAACCTTTGGTGTGCTTACAGTACCC 210  
Db 379 GAGCCTCGGGGAAGCACTTCAACGAGAGCCAGCATGTGCAAGTCTTCTCCCGG 438  
Qy 211 TACACCGCTGTCAAGAGACTTGTATGTTCCCTCCCTGCACACACCTCCCTGAGACATCCGTT 270  
Db 439 TTCACAGGTCAGGAGATTTGCAATGTTCCCTGATGTGCGGCCGCGGAGCACTA 498  
Qy 271 AGAGAGATTTGATGAGCTGGCTGATCTTCAACAGATGTGAAAGCTGCTGGAATGCAAG 330  
Db 499 GCCGAGATCGAGCTCTCTATGACGTCTACCTGATGTCCAAAGAAATGGGGCTTGGAG 558  
Qy 331 AATTTCAATTTTCATGAGTACTTCAATGCTGCTGCAAGTATGCTCCCAAGAGGCTTGG 390  
Db 559 GACGTCATGTTGATGGGGAATTCATATGCGGCTGACAGTATGTGAGAACCTCCAGTGG 618  
Qy 391 AAGGACATCCGCTGAGAGCGAGCCCAAGTGTGTTGGCTGATGCGGAGCAAGAGAC 450  
Db 619 TCATCCATCCGCTGTGAGCAAGGCCCACTTCCAGTGGCTGATCCCGACAGGCTGAC 678  
Qy 451 ACCAGGTCAGAGAGCAAACTGCGCTATGACAGATGTGCTTACAGAGCAAAAT 510



Db 679 ACCACAG---CTACACCCAGCAGCTGTGCTATGACAGGATCGTGTTCAGGGATGCTG 735  
 QY 511 ATTGTCAACTCTGGTGGTCTCAATAAACCCTCGTCTTTTGATTCCAGAAAGCTTACAGG 570  
 Db 736 CTCGAGGGGCGTGTTCGCGACTGGCTCTTCCCTTTAACTTCCAGGCTGCTATGSC 795  
 QY 571 TTGTCCTGAATCGAAGGCC 589  
 Db 796 CTGAGTGACCAACTGGCCC 814

RESULT 15

US-09-825-012-62  
 ; Sequence 62, Application US/09825012  
 ; Patent No. US20020122798A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Young, Robert  
 ; TITLE OF INVENTION: Compounds for Targeting  
 ; FILE REFERENCE: 43191-256808  
 ; CURRENT APPLICATION NUMBER: US/09/825,012  
 ; CURRENT FILING DATE: 2001-04-03  
 ; PRIOR APPLICATION NUMBER: US 60/237,159  
 ; PRIOR FILING DATE: 2000-10-02  
 ; PRIOR APPLICATION NUMBER: GB 0008049.9  
 ; PRIOR FILING DATE: 2000-04-03  
 ; NUMBER OF SEQ ID NOS: 102  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 62  
 ; LENGTH: 1548  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Humanised HMPG1 Fd - DNase I fusion  
 US-09-825-012-62

Query Match 25.0%; Score 147.8; DB 10; Length 1548;  
 Best Local Similarity 56.5%; Pred. No. 1.7e-36;  
 Matches 316; Conservative 0; Mismatches 237; Indels 6; Gaps 2;  
 QY 34 ACATACAACGTATTAGTCTCGCCTTGAAGAAACACATATAAAGACAGTATGCC 93  
 Db 949 ACCTATCACTAGTGTGCTAGTGAGCCACTGGGACGGAACAGCTATAAGGAGCGCTACCTG 1008  
 QY 94 TTTCTCTATAAAGAAAGCTAGTGTCTGTAAACAAAGCTACCTCTACCACGA---CTAT 150  
 Db 1009 TTGCGGTACAGCCCTGACCAGGTGCTCGGTGGACAGCTACTACTACGATGATGCTGC 1068  
 QY 151 CAGGCTGGAGACGAGATGTGTTTCCAGGGAAACCCTTTGTGGTCTGTGTTCCAGTCACCC 210  
 Db 1069 GAGCCCTGCGGGAACGACACCTTCAACCGAGAGCCAGCCATTGTCTAGGTTCTTCTCCCGG 1128  
 QY 211 TACACCGCTGTCAAGGACTTCGTGATGTGCCCTGCACACCCCTGAGACATCCGTT 270  
 Db 1129 TTCACAGAGTCAGGAGTTCGCCATTGTTCCTCCCTGCATGCGGCCCGGGGACGAGTA 1188  
 QY 271 AGAGAGATTGATGAGCTGCTGATGTCTACACAGATGTGAACGTCGCTGGAATGCAGAG 330  
 Db 1189 GCCGAGATCGACGCTCTCTATGACGCTACCTGGATGTCAGAGAAATGGGGCTTGGAG 1248  
 QY 331 AATTTCAATTTTCATGGGTGACTTCAATGTGCTGCTGAGCTACGTCCCCAAGAGGCTGG 390  
 Db 1249 GACGTCAATGTTGATGGGCACTTCAATGCGGCTGAGCTATGTGAGACCCCTCCAGTGG 1308  
 QY 391 AAGGACATCCGCTGAGGAGGACCCCAAGTTTCGTTGGCTGATCGGGACCAAGAGGAC 450  
 Db 1309 TCATCCATCCGCTGTGGACAAGCCCCACCTTCCAGTGGCTGATCCCCGACACGCTGAC 1368  
 QY 451 ACCACGGTCAAGAAGAGCACAACCTGCGCTATGACAGGATCGTGTCTAGAGACAAAAT 510  
 Db 1369 ACCACAG---CTACACCCACGCACTGTGCTATGACAGGATCGTGGTTCAGGGATGCTG 1425  
 QY 511 ATGTCAACTCTGGTGGTCTCAATAAACCTCGTCTTTTGATTTCCAGAAAGCTTACAGG 570

Db 1426 CTCGAGGGGCGTGTTCGCCGACTCGGCTCTTCCCTTTAACTTCCAGGCTGCTATGSC 1485  
 QY 571 TTGTCCTGAATCGAAGGCC 589  
 Db 1486 CTGAGTGACCAACTGGCCC 1504  
 Search completed: December 9, 2002, 09:24:07  
 Job time : 127.65 secs

**This Page Blank (uspto)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 09:18:16 ; Search time 1408.41 Seconds  
(without alignments)  
13865.275 Million cell updates/sec

Title: US-09-905-114-3  
Perfect score: 671  
Sequence: 1 acaacagatgtgccccata.....atcatcatgaagaaccatga 671

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	527	78.5	1023	6	AX329742	AX329742 Sequence
2	527	78.5	1023	6	AX332286	AX332286 Sequence
3	527	78.5	1023	6	AX332789	AX332789 Sequence
4	527	78.5	1023	6	AX410746	AX410746 Sequence
5	527	78.5	1023	6	HSU56814	U56814 Human DNase
6	527	78.5	1067	9	BC015831	BC015831 Homo sapi
7	527	78.5	1079	9	AF047354	AF047354 Homo sapi
8	525.4	78.3	1108	9	HSU75744	U75744 Homo sapien
9	501.4	74.7	1331	10	AF039852	AF039852 Rattus norv
10	501.4	74.7	1417	10	RNU75689	U75689 Rattus norv
11	481.2	71.7	1208	6	AR047845	AR047845 Sequence
12	481.2	71.7	1208	6	E11687	E11687 cDNA encodi
13	474.8	70.8	941	10	MMU76110	U76110 Mus musculu
14	474.8	70.8	1124	10	AF047355	AF047355 Mus muscu
15	473.2	70.5	2110	10	BC012671	BC012671 Mus muscu
16	411.4	61.3	1223	6	AX400016	AX400016 Sequence
17	298.4	44.5	1250	5	AF059612	AF059612 Xenopus l
18	195.6	29.2	392	6	AX330074	AX330074 Sequence
19	195.6	29.2	392	6	AX407983	AX407983 Sequence
20	182.4	27.2	1295	4	BTPANDNAI	AJ001538 Bos tauru
21	181.4	27.0	868	4	AF311922	AF311922 Bos tauru
22	173.8	25.9	1283	5	AB041732	AB041732 Cynops py
23	172.6	25.7	1210	5	AB038776	AB038776 Rana cate
24	164.8	24.6	1161	4	D82875	D82875 Oryctolagus
25	163.2	24.3	1039	6	AX030575	AX030575 Sequence
26	163.2	24.3	1039	6	AX030591	AX030591 Sequence
27	163.2	24.3	1039	6	AX268671	AX268671 Sequence
28	163.2	24.3	1039	9	HUMDNASEI	M55983 Human DNase
29	163.2	24.3	1149	9	BC029437	BC029437 Homo sapi
30	163.2	24.3	2054	9	HSA298844	AJ298844 Homo sapi
31	161.6	24.1	783	6	AX268673	AX268673 Sequence
32	161.6	24.1	858	6	AX268675	AX268675 Sequence
33	161.6	24.1	1039	6	AR019468	AR019468 Sequence
34	161.6	24.1	1312	10	BC014718	BC014718 Mus muscu
35	161.6	24.1	1337	10	BC030394	BC030394 Mus muscu
36	161.6	24.1	1548	6	AX268721	AX268721 Sequence
37	161.6	24.1	1548	6	AX268722	AX268722 Sequence
38	161.6	24.1	1554	6	AX268701	AX268701 Sequence
39	161.6	24.1	1554	6	AX268702	AX268702 Sequence
40	161.6	24.1	1557	6	AX268723	AX268723 Sequence
41	161.6	24.1	1560	6	AX268729	AX268729 Sequence
42	161.6	24.1	1560	6	AX268730	AX268730 Sequence
43	161.6	24.1	1560	6	AX268733	AX268733 Sequence
44	161.6	24.1	1560	6	AX268734	AX268734 Sequence
45	161.6	24.1	1563	6	AX268703	AX268703 Sequence

ALIGNMENTS

RESULT 1  
AX329742  
LOCUS AX329742  
DEFINITION Sequence 251 from Patent WO0194629.  
ACCESSION AX329742  
VERSION AX329742.1 GI:18102720  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,  
Horrigan, S., Soppet, D.R. and Weaver, Z.  
Cancer gene determination and therapeutic screening using signature

Gene sets  
Patent: WO 0194629-A 251.13-DEC-2001;  
Avalon Pharmaceuticals (US)

FEATURES  
source  
1.1023  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 312 a 244 c 240 g 227 t  
ORIGIN

Query Match 78.5%; Score 527; DB 6; Length 1023;  
Best Local Similarity 87.8%; Pred. No. 1.1e-132;  
Matches 575; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 1 ACAACAGGATTCGCCCTACTGATGAGAGAGCTTAAACGAAATTCAGAAAGGCAATTA 60  
DB 215 ACAACAGGATTCGCCCTACTGATGAGAGAGCTTAAACGAAATTCAGAAAGGCAATTA 274  
QY 61 CATACACTATGTGATTAAGCTCTGCTGCTGGAAGAAACATATTAAGACATATGCTT 120  
DB 275 CGTCAACTATGTGATTAAGCTCTGCTGCTGGAAGAAACATATTAAGACATATGCTT 334  
QY 121 TTCTCTATAAGAAAGCTAGTGTCTGTAAACAAAGCTTCTTACACGACTATCAAG 180  
DB 335 TTCTCTATAAGAAAGCTAGTGTCTGTAAACAAAGCTTCTTACACGACTATCAAG 394  
QY 181 CTGAGAGCGAGATGTGTTTCCAGGAAACCTTTGTGCTGTGTTCCAGTCACTTCA 240  
DB 395 ATGGAGAGCGAGATGTGTTTCCAGGAAACCTTTGTGCTGTGTTCCAGTCACTTCA 454  
QY 241 CCGCTGTCAAGAGCTTCTGATTTGCTCCCTGCAACACCCCTGAGACATCCGTTAAG 300  
DB 455 CTGCTGTCAAGAGCTTCTGATTTGCTCCCTGCAACACCCCTGAGACATCCGTTAAG 514  
QY 301 AGATTGATGAGCTGCTGATGTTTACACAGATGTAAACGCTGTGAATGCAAGATT 360  
DB 515 AGATTGATGAGCTGCTGATGTTTACACAGATGTAAACGCTGTGAATGCAAGATT 574  
QY 361 TCATTTTTCATGAGTGTGATCTTCAATGCTGCTGAGTACCTCCCAAGAGGCTTGAAG 420  
DB 575 TCATTTTTCATGAGTGTGATCTTCAATGCTGCTGAGTACCTCCCAAGAGGCTTGAAG 634  
QY 421 ACATCCGCTGAGAGAGAGACCCCAAGTTGCTTGGCTGATCGGGAGCCAAAGAGACCA 480  
DB 635 ACATCCGCTGAGAGAGAGACCCCAAGTTGCTTGGCTGATCGGGAGCCAAAGAGACCA 694  
QY 481 CGGTCAAGAGAGAGACCAAACTGCGCTTATGACAGAGATCGTGTGAAGAGACAAATATTG 540  
DB 695 CGGTCAAGAGAGAGACCAAACTGCGCTTATGACAGAGATCGTGTGAAGAGACAAATATTG 754  
QY 541 TCAACTCTGTGAGTGTCTCAATCAAACTGCTTGAATTTCCAGAAAGCTTACAGTTGT 600  
DB 755 TCAACTCTGTGAGTGTCTCAATCAAACTGCTTGAATTTCCAGAAAGCTTACAGTTGT 814  
QY 601 CTGAATCGAAGGCGCTGATGTCAAGGACCACTTTCCAGTTCAATCATCATCATCA 655  
DB 815 CTGAATCGAAGGCGCTGATGTCAAGGACCACTTTCCAGTTCAATCATCATCATCA 869

RESULT 2  
AX332286 1023 bp DNA linear PAT 09-JAN-2002  
LOCUS AX332286  
DEFINITION Sequence 2795 from Patent WO0194629.  
ACCESSION AX332286  
VERSION AX332286.1 GI:18122920  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
1 Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endress, G.,  
Horrikan, S., Sopet, D. R. and Weaver, Z.

TITLE Cancer gene determination and therapeutic screening using signature

gene sets  
Patent: WO 0194629-A 2795.13-DEC-2001;  
Avalon Pharmaceuticals (US)

FEATURES  
source  
1.1023  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 312 a 244 c 240 g 227 t  
ORIGIN

Query Match 78.5%; Score 527; DB 6; Length 1023;  
Best Local Similarity 87.8%; Pred. No. 1.1e-132;  
Matches 575; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 1 ACAACAGGATTCGCCCTACTGATGAGAGAGCTTAAACGAAATTCAGAAAGGCAATTA 60  
DB 215 ACAACAGGATTCGCCCTACTGATGAGAGAGCTTAAACGAAATTCAGAAAGGCAATTA 274  
QY 61 CATACACTATGTGATTAAGCTCTGCTGCTGGAAGAAACATATTAAGACATATGCTT 120  
DB 275 CGTCAACTATGTGATTAAGCTCTGCTGCTGGAAGAAACATATTAAGACATATGCTT 334  
QY 121 TTCTCTATAAGAAAGCTAGTGTCTGTAAACAAAGCTTCTTACACGACTATCAAG 180  
DB 335 TTCTCTATAAGAAAGCTAGTGTCTGTAAACAAAGCTTCTTACACGACTATCAAG 394  
QY 181 CTGAGAGCGAGATGTGTTTCCAGGAAACCTTTGTGCTGTGTTCCAGTCACTTCA 240  
DB 395 ATGGAGAGCGAGATGTGTTTCCAGGAAACCTTTGTGCTGTGTTCCAGTCACTTCA 454  
QY 241 CCGCTGTCAAGAGCTTCTGATTTGCTCCCTGCAACACCCCTGAGACATCCGTTAAG 300  
DB 455 CTGCTGTCAAGAGCTTCTGATTTGCTCCCTGCAACACCCCTGAGACATCCGTTAAG 514  
QY 301 AGATTGATGAGCTGCTGATGTTTACACAGATGTAAACGCTGTGAATGCAAGATT 360  
DB 515 AGATTGATGAGCTGCTGATGTTTACACAGATGTAAACGCTGTGAATGCAAGATT 574  
QY 361 TCATTTTTCATGAGTGTGATCTTCAATGCTGCTGAGTACCTCCCAAGAGGCTTGAAG 420  
DB 575 TCATTTTTCATGAGTGTGATCTTCAATGCTGCTGAGTACCTCCCAAGAGGCTTGAAG 634  
QY 421 ACATCCGCTGAGAGAGAGACCCCAAGTTGCTTGGCTGATCGGGAGCCAAAGAGACCA 480  
DB 635 ACATCCGCTGAGAGAGAGACCCCAAGTTGCTTGGCTGATCGGGAGCCAAAGAGACCA 694  
QY 481 CGGTCAAGAGAGAGACCAAACTGCGCTTATGACAGAGATCGTGTGAAGAGACAAATATTG 540  
DB 695 CGGTCAAGAGAGAGACCAAACTGCGCTTATGACAGAGATCGTGTGAAGAGACAAATATTG 754  
QY 541 TCAACTCTGTGAGTGTCTCAATCAAACTGCTTGAATTTCCAGAAAGCTTACAGTTGT 600  
DB 755 TCAACTCTGTGAGTGTCTCAATCAAACTGCTTGAATTTCCAGAAAGCTTACAGTTGT 814  
QY 601 CTGAATCGAAGGCGCTGATGTCAAGGACCACTTTCCAGTTCAATCATCATCATCA 655  
DB 815 CTGAATCGAAGGCGCTGATGTCAAGGACCACTTTCCAGTTCAATCATCATCATCA 869

RESULT 3  
AX332789 1023 bp DNA linear PAT 10-JAN-2002  
LOCUS AX332789  
DEFINITION Sequence 3298 from Patent WO0194629.  
ACCESSION AX332789  
VERSION AX332789.1 GI:18123423  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
1 Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endress, G.,

TITLE	Horrigan, S., Soppet, D.R. and Weaver, Z. Cancer gene determination and therapeutic screening using signature gene sets			
JOURNAL	Patent: WO 0194629-A 3298 13-DEC-2001; Avalon Pharmaceuticals (US)			
FEATURES	Location/Qualifiers 1..1023			
source	/organism="Homo sapiens" /db_xref="taxon:9606"			
BASE COUNT	312 a	244 c	240 g	227 t
ORIGIN				

Query Match	78.5%;	Score 527;	DB 6;	Length 1023;
Best Local Similarity	87.8%;	Pred. No. 1.1e-132;		
Matches 575;	Conservative	0;	Mismatches 80;	Indels 0; Gaps 0;
1	ACAAAGGATCTGCCCATCTATGATGGAGAGCTAAACGGAATTTCAAGAAAGGCATAA	60		
215	ACAAAGGATCTGCCCATCTATGATGGAGAGCTAAACGGAATTTCAAGGAGAGGCATAA	274		
61	CATACAACTATGTGATTAGCTCTCGCCTTTGGAAGAAACACATATAAAGAACAGTATGCCT	120		
275	CGTACAACTATGTGATTAGCTCTCGCCTTTGGAAGAAACACATATAAAGAACATATGCCT	334		
121	TTCTCTATAAAGAAAGCTAGTGTCTGTAAACAAAGCTACTCTACACGACTATCAGG	180		
335	TTCTCTACAAAGAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG	394		
181	CTGGAGACGCAGATGTGTTTTCCAGGAAACCTTTTGTGGTCTGGTTCCAGTCAACCTTACA	240		
395	ATGGAGACGCAGATGTGTTTTCCAGGAGGCCCTTTTGTGGTCTGGTTCCAAATCTCCCCACA	454		
241	CGCGTGTCAAGGACATTCGTGATTGTCCCTCCACACACCCCTCAGACATCCCGTTAGAG	300		
455	CTGCTGTCAAAGACATTCGTGATTATCCCTCCACACACCCACGAGACATCCCGTTAAGG	514		
301	AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAACCGTCCGTGGAAATGCAGAGAATT	360		
515	AGATCGATGAGTTGGTTGAGGTTCTACCGGACGTGAACACCGCTGGAAAGCGGGAATT	574		
361	TCATTTTCATGGGTGACTTTCAATGTGGCTGAGCTACGTCCCCAAGAGGCCTGGAAAG	420		
575	TCATTTTCATGGGTGACTTTCAATGTGGCTGAGCTACGTCCCCAAGAGGCCTGGAAAG	634		
421	ACATCCGCTGAGGACGGACCCCAAGTTCTGTTGGCTGATCGGGACCAAGAGGACCA	480		
635	ACATCCGCTGAGGACGTACCCCAAGTTCTGTTGGCTGATCGGGACCAAGAGGACCA	694		
481	CGGTCAAGAAAGACACAAACTCGCCTATGACAGGATCGTCTTAGAGGACAAAAATATTG	540		
695	CGGTGAAGAAAGACACAAACTGTGCATATGACAGATTGTGCTTAGAGGACAAAGAAATCG	754		
541	TCAACTCTCGTGGTCTCCTCAATCAAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGTTGT	600		
755	TCAGTTCTGTGTTGCCCAAGTCAACACAGTGTGTTTGTGCTTCCAGAAAGCTTCAAGCTGA	814		
601	CTGAATCGAAGGCCCTGGATGTGACGGACCACTTTCCAGTTTCATCATCATCATCA	655		
815	CTGAAGAGAGGAGCCCTGGATGTGACGGACCACTTTCCAGTTTGAATTTTAACTATCA	869		

RESULT 4	AX410746	1023 bp	DNA	linear	PAT 14-JUN-2002
LOCUS	AX410746	Sequence 3393	from Patent WO0229103.		
DEFINITION	AX410746				
ACCESSION	AX410746				
VERSION	AX410746.1	GI:21443451			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1				

AUTHORS Alvares C., Horne D. Peres-da-Silva S. and Vockley J.C.  
 TITLE Gene expression profiles in liver cancer  
 JOURNAL Patent: WO 0229103-A 3393 11-APR-2002;  
 GENE LOGIC INC. (US)  
 FEATURES Location/Qualifiers  
     source 1..1023  
         /organism="Homo sapiens"  
         /db\_xref="taxon:9606"  
         /notes="EMBL/GenBank Accession No. U56814"  
 BASE COUNT 312 a 244 c 240 g 227 t  
 ORIGIN

Query Match	78.5%;	Score 527;	DB 6;	Length 1023;
Best Local Similarity	87.8%;	Pred. No. 1.1e-132;		
Matches 575;	Conservative	0;	Mismatches 80;	Indels 0; Gaps
Qy	1	ACACAGGATCTGCCCATACTGATGGAGAAGCTAAACGGAAATTCAACAAAGGCATAA	60	
Db	215	ACACAGGATCTGCCCATACTGATGGAGAAGCTGAACAGAANTTCAGGAGGCGCATAA	274	
Qy	61	CATACAACTATGTGATTAGTCTTCGCCTTGGAGAAACACATATAAGAAACAGTATGCCT	120	
Db	275	CGTACACTATGTGATTAGTCTTCGCCTTGGAGAAACACATATAAGAAACAAATATGCCT	334	
Qy	121	TTCTCTATAAGAAAAAGCTAGTGTCTGTATAAACAAAGCTACCTCTACACAGACTATCAGG	180	
Db	335	TTCTCTACAAAGAAAAGCTGGTGTCTGTGAAGAGGAGTTATCACTACCATGACTATCAGG	394	
Qy	181	CTGGAGACGAGATGTGTTTTCAGGGAACCTTTGTGTCTGTGTCTGCAGTACCCCTACA	240	
Db	395	ATGGAGACGAGATGTGTTTTCAGGAGGAGCCCTTGTGTGTCTGGTTCCTCAATCTCCCCACA	454	
Qy	241	CCGCTGTCAAGGACTTCGTGATTGTCCCTCTGCACACACCCCTGAGAGCATCCGTTTAGAG	300	
Db	455	CTGCTGTCAAGAGACTTCGTGATTATCCCCCTGCACACACCCACAGAGACATCCGTTTAGG	514	
Qy	301	AGATTGTAGCTGGGTGATGTCCTACACAGATGTGAAACGTCGCTGGAAATGCAGAGAATT	360	
Db	515	AGATCGATGAGTTGGTTGAGGTCTACACGGACGTGAAACACCGCTGGAAGGCGGAGAATT	574	
Qy	361	TCATTTTTCATGGGTGACTTCAATGCTGGGTGAGCTACGTCCTCCCAAGAAAGGCTTGAAGG	420	
Db	575	TCATTTTTCATGGGTGACTTCAATGCGCGGTTCGACGCTACGCTCCCAAGAAAGGCTTGAAGA	634	
Qy	421	ACATCCGCTTGAGGACGGACCCCAAGTCGTTGGCTGATCGGGGACCAAGAGGACACCA	480	
Db	635	ACATCCGCTTGAGGACTGACCCCAAGTTGTTTGGCTGATCGGGGACCAAGAGGACACCA	694	
Qy	481	CGGTCAAGAAGAGCACAAACTCGGCGCTATGACAGGATCGTGCTTAGAGGACAAAATATTG	540	
Db	695	CGGTGAAGAAGAGCACCAACTGTGTCATATGACAGGATTTGCTTAGAGGACCAAGAAATCG	754	
Qy	541	TCAACTCTGGTGGTCTCAATCAAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT	600	
Db	755	TCAGTTCTGTTGTTCCCAAGTCAACAGTGTTTTTCGACTTCCAGAAAGCTTACAAGCTGA	814	
Qy	601	CTGAATCGAAGGCCCTCGGATGTGAGGACCACTTTCAGTTTCATCATCATCATCA	655	
Db	815	CTGAAGAGGAGGCCCTCGGATGTGAGGACCACTTTCAGTTGAATTTAACTACATA	869	

RESULT 5  
HSU56814  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Eukaryota; Metazoa; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1023)  
1023 bp mRNA linear PRI 24-JUL-1997  
Human Dnaasei-Like III protein (DNAS1L3) mRNA, complete cds.

AUTHORS Rodriguez, A.M., Rodin, D., Nomura, H., Morton, C.C., Weremowicz, S. and Schneider, M.C.  
 TITLE Identification, localization, and expression of two novel human genes similar to deoxyribonuclease I  
 JOURNAL Genomics 42 (3), 507-513 (1997)  
 MEDLINE 97349121  
 PUBMED 9205125  
 REFERENCE 2 (bases 1 to 1023)  
 AUTHORS Schneider, M.C. and Rodriguez, A.  
 JOURNAL Direct Submission  
 Submitted (25-APR-1996) M.C. Schneider, Renal Division, Brigham and Women's Hospital, 75 Francis Street, Boston, MA 02115, USA  
 FEATURES  
 Source  
 1. 1023  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="Stratagene liver (#937224)"  
 /note="Corresponds to EST clones 82269 (Genbank Accession Number T66985, T69063), 82738 (Genbank Accession Number T73558, T73653), 78422 (Genbank Accession Number T61400, T61368)"  
 1. 1023  
 /gene="DNAS1L3"  
 25..942  
 /gene="DNAS1L3"  
 /codon\_start=1  
 /product="DNaseI-Like III protein"  
 /protein\_id="AAB63967.1"  
 /db\_xref="GI:1399719"  
 /translation="MSRELIAPLLLLLSIALMRICSPNVRSFGSKQEDKNAMV IVKVIKCDIILVMEIKDSNNRCPIIMEKLNKSRGITNYVYSRLGRNTYKEY AFLYKEKLVSVKRSYHYHDYQDDADVPFSRPFVWFQSHTVAKDFVILPLHTTPT SVKSIDLEVEYTDVHKRWAKENFIIMGDFNAGSVYPKAKMKNIRLTDPRFVWLIG DQEDTTVKSTNCAYDRIVLRCGEIVSVVPSKNSVDFQKAYKLTLEEDLVSDHFP VEFKLOSRAFTNSKSVTLRKKTSKRS"  
 BASE COUNT 312 a. 244 c 240 g 227 t  
 ORIGIN  
 Query Match 78.5%; Score 527; DB 9; Length 1023;  
 Best Local Similarity 87.8%; Pred. No. 1,1e-132;  
 Matches 575; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

Qy 481 CGGTCAAGAGAGCAAACTGCCCTATGACAGATGCTGCTTAGAGGACAAAATATTG 540  
 Db 695 CGGTGAAGAAAGAGACCAACTGTGCATATGACAGATTTGCTTAGAGGACAAAATTCG 754  
 Qy 541 TCAACTCGTGTGTCCTCAATCAACCTCGCTTGAATTCACAAAGCTTAGAGTTGT 600  
 Db 755 TCAGTCTGTTGTTCCCAAGTCAACAGTGTGTTTGACCTTCAGAAAGCTTAGAGTGA 814  
 Qy 601 CTGAATCGAAGCCCTGGATGTCAGGACCACTTTCAGTTCATCATCATCATCA 655  
 Db 815 CTGAAGAGAGAGCCCTGGATGTCAGGACCACTTTCAGTTCATCATCATCA 869  
 RESULT 6  
 BC015831  
 LOCUS  
 DEFINITION Homo sapiens, deoxyribonuclease I-like 3, clone MGC:27146  
 IMAGE:4723725, mRNA, complete cds.  
 ACCESSION BC015831  
 VERSION BC015831.1 GI:16198370  
 KEYWORDS MGC.  
 SOURCE  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1067)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
 REMARK  
 COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [gcaps-re@mail.nih.gov](mailto:gcaps-re@mail.nih.gov)  
 Tissue Procurement: CLONTECH  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLI)  
 DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
 contact: [amad@systemsbiology.org](mailto:amad@systemsbiology.org)  
 Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia Greene, Mark Ketterman and Anuradha Madan  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLI at: <http://image.llnl.gov>  
 Series: IRAL Plate: 38 Row: P Column: 7  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4826697.  
 FEATURES  
 source  
 1. 1067  
 /organism="Homo sapiens"  
 /db\_xref="LocusID:1776"  
 /db\_xref="taxon:9606"  
 /clone="MGC:27146 IMAGE:4723725"  
 /tissue\_type="Liver"  
 /clone\_lib="NIH MGC\_76"  
 /lab\_host="DH10B"  
 /note="Vector: PDNR-LIB"  
 68..985  
 /codon\_start=1  
 /product="deoxyribonuclease I-like 3"  
 /protein\_id="AAH15831.1"  
 /db\_xref="GI:16198371"  
 /translation="MSRELIAPLLLLLSIALMRICSPNVRSFGSKQEDKNAMV IVKVIKCDIILVMEIKDSNNRCPIIMEKLNKSRGITNYVYSRLGRNTYKEY AFLYKEKLVSVKRSYHYHDYQDDADVPFSRPFVWFQSHTVAKDFVILPLHTTPT SVKSIDLEVEYTDVHKRWAKENFIIMGDFNAGSVYPKAKMKNIRLTDPRFVWLIG DQEDTTVKSTNCAYDRIVLRCGEIVSVVPSKNSVDFQKAYKLTLEEDLVSDHFP VEFKLOSRAFTNSKSVTLRKKTSKRS"  
 BASE COUNT 327 a 256 c 249 g 235 t  
 ORIGIN

Query Match		78.5%; Score 527; DB 9; Length 1067;
Best Local Similarity		87.8%; Pred. No. 1.1e-132;
Matches 575; Conservative		0; Mismatches 80; Indels 0; Gaps 0;
QY	1	ACACAGGATCTGCCCATCTACTGATGAGAGCTAAACCGAAATTCAGAAAAGGCATAA 60
	258	ACACAGGATCTGCCCATCTACTGATGAGAGCTGACAGAAATTCAGAGAGGCATAA 317
QY	61	CATACAACTATGTGATTAGCTCTCGCTTGGAGAAACACATATAAGAACATGATGCCT 120
	318	CGTAACTATGTGATTAGCTCTCGCTTGGAGAAACACATATAAGAACATATGCTT 377
QY	121	TTCTCTATAAGAAAGCTAGTGTCTGTAAACAAAGCTACCTTACCAGCATATCAGG 180
	378	TTCTCTACAAGAAAGCTGGTGTCTGTGAAGAGGAGTTTACATACCATGATATCAGG 437
QY	181	CTGGAGACGAGATGTGTTTCCAGGAAACCTTTGTGCTGCTGCTGCTGCTGCTGCT 240
	438	ATGGAGACGAGATGTGTTTCCAGGAGCCCTTTGTGCTGCTGCTGCTGCTGCTGCT 497
QY	241	CGCTCTCAAGGACTTCGTGATTGCTCCCTGCACACACCCCTGAGACATCGTTAGAG 300
	498	CTGCTGTCAAGACTTCGTGATTGCTCCCTGCACACACCCCTGAGACATCGTTAGAG 557
QY	301	AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAACGCTCGCTGGAATGCAGAAAT 360
	558	AGATCGATGAGTTGGTTGAGGTCTACACGAGCTGAAACACCGCTGGAAGCGGAGAA 617
QY	361	TCAATTTTCAATGAGTCTCAATGCTGGCTGACGCTACGCTCCCAAGAGGCTGGAAG 420
	618	TCAATTTTCAATGAGTCTCAATGCTGGCTGACGCTACGCTCCCAAGAGGCTGGAAG 677
QY	421	ACATCGCTGAGGACCGACCCCAAGTTCGTTGGCTGATCGGGACCAAGAGCACCA 480
	678	ACATCGCTGAGGACCGACCCCAAGTTCGTTGGCTGATCGGGACCAAGAGCACCA 737
QY	481	CGGTCAAGAGAGCAAACTGCGCTATGACAGGATCGTGTCTAGAGGACAAATATTG 540
	738	CGGTGAAGAGAGCAAACTGCGCTATGACAGGATCGTGTCTAGAGGACAAAGAAATCG 797
QY	541	TCACTCTGGTGTCTCAATCAAACTCGTCTTTGATTTCCAGAAAGCTTACAGTTGT 600
	798	TCACTCTGGTGTCTCAATCAAACTCGTCTTTGATTTCCAGAAAGCTTACAGTTGT 857
QY	601	CTGAATCGAAGCCCTGGATGTGACGACCACTTCCAGTTCATCATCATCA 655
	858	CTGAAGAGGAGCCCTGGATGTGACGACCACTTCCAGTTCATCATCATCA 912
RESULT 7		
LOCUS AF047354		
DEFINITION Homo sapiens liver and spleen DNase precursor (LSD) mRNA, complete cds.		
ACCESSION AF047354		
VERSION AF047354.1 GI:2905785		
KEYWORDS		
SOURCE Homo sapiens.		
ORGANISM Homo sapiens		
REFERENCE 1 (bases 1 to 1079)		
AUTHORS Baron, W.F., Pan, C.Q., Spencer, S.A., Ryan, A.M., Lazarus, R.A. and Baker, K.P.		
TITLE Cloning and characterization of an actin-resistant DNase I-like endonuclease secreted by macrophages		
JOURNAL Gene 215 (2), 291-301 (1998)		
MEDLINE 98382522		
PUBMED 9714828		
REFERENCE 2 (bases 1 to 1079)		
AUTHORS Baron, W.F., Pan, C.Q., Spencer, S.A., Ryan, A.M., Lazarus, R.A. and Baker, K.P.		
TITLE Direct Submission		
JOURNAL Submitted (09-FEB-1998) Molecular Biology, Genentech Inc., 1 DNA Way, South San Francisco, CA 94080, USA		
FEATURES	Location/Qualifiers	
	source	1..1079
gene	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
CDS	/gene="LSD"	
	/note="actin resistant activity; macrophage specific DNase I; LS-DNase"	
sig_peptide	/codon_start=1	
	/product="liver and spleen DNase precursor"	
mat_peptide	/protein_id="AAC35752.1"	
	/db_xref="GI:2905786"	
BASE COUNT	330 a 260 c 252 g 237 t	
ORIGIN	/evidence=experimental	
	/gene="LSD"	
Query Match	78.5%; Score 527; DB 9; Length 1079;	
	Best Local Similarity 87.8%; Pred. No. 1.1e-132;	
QY	1	ACAACAGGATCTGCCCATCTACTGATGAGAGCTAAACCGAAATTCAGAAAAGGCATAA 60
	261	ACAACAGGATCTGCCCATCTACTGATGAGAGCTGAAACAGAAATTCAGGAGAGCATAA 320
QY	61	CATACAACTATGTGATTAGCTCTCGCTTGGAGAAACACATATAAGAACATGATGCCT 120
	321	CGTAACTATGTGATTAGCTCTCGCTTGGAGAAACACATATAAGAACATATGCTT 380
QY	121	TTCTCTATAAGAAAGCTAGTGTCTGTAAACAAAGCTACCTTACCAGCATATCAGG 180
	381	TTCTCTACAAGAAAGCTGGTGTCTGTGAAGAGGAGTTTACATACCATGATCATCAG 440
QY	181	CTGGAGACGAGATGTGTTTCCAGGAAACCCCTTTGTGCTGCTGCTGCTGCTGCTGCT 240
	441	ATGGAGACGAGATGTGTTTCCAGGAGCCCTTTGTGCTGCTGCTGCTGCTGCTGCTGCT 500
QY	241	CGCTGTCAAGAGACTTGTGATTGTCCTCCCTGACACACCCCTGAGACATCCGTTAGAG 300
	501	CTGCTGTCAAGAGACTTGTGATTGTCCTCCCTGACACACCCCTGAGACATCCGTTAGAG 560
QY	301	AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAACGCTCGCTGGAATGCAGAGAATT 360
	561	AGATCGATGAGTTGGTTGAGGTCTACCGAGCTGGAACACCCCTGGAAGCGCGGAATT 620
QY	361	TCATTTTCATGGTGACTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
	621	TCATTTTCATGGTGACTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 680
QY	421	ACATCCGCTGAGGACGGACCCCAAGTTGCTTGGCTGATCGGGACCAAGAGGACCA 480
	681	ACATCCGCTGAGGACGGACCCCAAGTTGCTTGGCTGATCGGGACCAAGAGGACCA 740
QY	481	CGGTCAAGAGAGCAAACTGCGCTATGACAGGATCGTGTCTAGAGGACAAATATTG 540
	741	CGGTGAAGAGAGCAAACTGCGCTATGACAGGATCGTGTCTAGAGGACAAAGAAATCG 800
QY	541	TCAACTCTGGTGTCTCAATCAAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGTTGT 600



```

Db      801  TCAGTTCTGTTTCCCAAGTCAACAGTGTTCAGTTCAGAAAGTTTCAAGCTGA 860
Oy      601  CTGAATCGAAGGCCCTGATGTGACGACCACTTCCAGTTCATCATCATCA 655
Db      861  CTGAAGAGAGAGCCCTGATGTGACGACCACTTCCAGTTCATTAACACTACA 915

RESULT 8
LOCUS   HSU75744 1108 bp mRNA linear PRI 17-JUN-1998.
DEFINITION Homo sapiens DNase gamma mRNA, complete cds.
ACCESSION U75744
VERSION   U75744.1 GI:3236319
KEYWORDS
SOURCE    Homo sapiens.
ORGANISM  Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;
            Eukaryota; Metazoa; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1108)
            Shiohawa, D., Hirai, M. and Tanuma, S.
            CDNA cloning of human DNase gamma: chromosomal localization of its
            gene and enzymatic properties of recombinant protein
            Apoptosis 3 (2): 89-95 (1998)
REFERENCE 2 (bases 1 to 1108)
            Shiohawa, D. and Tanuma, S.
            Direct Submission
            Submitted (23-OCT-1996) Biochemistry, Science Univ. of Tokyo,
            Shinjuku-Ku Ichigaya Funagawaracho, Tokyo 162, Japan
FEATURES
    source
        1..1108
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /function="apoptosis"
            /function="deoxyribonuclease"
            /codon_start=1
            /product="DNase gamma"
            /protein_id="AAC23652.1"
            /db_xref="GI:3236320"
            /translation="MSRELAPLLLLLSIHSLAMRISFNVRSPGSKOEDKXAMDV
            IVKIRCDIILWEIKDSNNRIPILMEKINRNRSGITNYVVISRLGKTKEDY
            AFLYKELVSKSYHHYDGDADVSRFPVWQSPHTANKDVIPLHTTPT
            SVKEIDELVEYTDVVKHKAENFIEMGDENAGCSYVFKKMKXIRLRITDRFVWLIG
            DQEDTIVKSTNCAYDRIVLAGEIVSVVSKSVDFQKAYKLTEEDALDVSDHFP
            VEFKLOSRAFTNSKSVTLKTKTKRSKS"
BASE COUNT 329 a 261 c 271 g 247 t
ORIGIN
Query Match 78.3%; Score 525.4; DB 9; Length 1108;
Beet Local Similarity 87.6%; Pred. No. 3,1e-132;
Matches 574; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

Oy      1  ACAAGAGATCGCCCACTACTGATGAGAGCTAAACGAATTCAGAAAGGCATPA 60
Db      283  ACAAGAGATCGCCCACTACTGATGAGAGCTAAACGAATTCAGAGAGGCATPA 342
Oy      61  CATACAACTATGTAGTATGAGCTTCGCTTGAAGAAACAATATAAGAACATAGTGCCT 120
Db      343  CGTACAACTATGTAGTATGAGCTTCGCTTGAAGAAACAATATAAGAACATATGCTT 402
Oy      121  TTCTTATAAAGAAAGCTAGTGTCTGTAACAAAGCTACTTACCAAGCATCTACG 180
Db      403  TTCTTATAAAGAAAGCTAGTGTCTGTAAGAGAGATGATCATACATGACTATCAG 462
Oy      181  CTGAGAGCAGATGTGTTTCCAGGAACCTTGTGTGCTGTTCCAGTCAACCTTACA 240
Db      463  ATGGAGAGCAGATGTGTTTCCAGGAACCTTGTGTGCTGTTCCAGTTCCTCCACA 522
Oy      241  CCGCTGTCAAGACTTCGTATGTGCTCCCTGACACACACCCCTGAGACATCCGTTAAG 300
Db      523  CTGCTGTCAAGACTTCGTATGTGCTCCCTGACACACACCCCTGAGACATCCGTTAAG 582

```

```

Oy      301  AGATTGATGACCTGCTGATGCTTACACAGATGTGAAACGTGCTGGAATGCAGAAATT 360
Db      583  AGATGATGATGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 642
Oy      361  TCATTTCATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
Db      643  TCATTTCATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 702
Oy      421  ACATCCGCTGAGAGACGACCCCAAGTTGTTGGTGTGATGCGGAGACCAAGAGACCA 480
Db      703  ACATCCGCTGAGAGACGACCCCAAGTTGTTGGTGTGATGCGGAGACCAAGAGACCA 762
Oy      481  CGGTCAAGAGACCAAACTGCGCTTATGACAGAGATGTCCTTAGAGACCAAAATATTG 540
Db      763  CGGTCAAGAGACCAAACTGCGCTTATGACAGAGATGTCCTTAGAGACCAAAATATTG 822
Oy      541  TCAACTCGTGTGATGCTTCAATCAAACTCGTCTTGTGATTCAGAAAGCTTACAGTTGT 600
Db      823  TCAGTTCTGTTTCCCAAGTCAACAGTGTGTTTGAATTCAGAAAGCTTACAGTTGT 882
Oy      601  CTGAATCGAAGGCCCTGATGTGACGACCACTTCCAGTTCATCATCATCATCA 655
Db      883  CTGAAGAGAGAGCCCTGATGTGACGACCACTTCCAGTTCATTAACACTACA 937

RESULT 9
LOCUS   AF039852 1331 bp mRNA linear ROD 07-AUG-1998
DEFINITION Rattus norvegicus DNasey mRNA, complete cds.
ACCESSION AF039852
VERSION   AF039852.1 GI:3395771
KEYWORDS
SOURCE    Rattus norvegicus.
ORGANISM  Rattus norvegicus; Chordata; Craniata; Vertebrata; Euteleostomi;
            Eukaryota; Metazoa; Primates; Sciuromorphi; Muridae; Murinae;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
            1 (bases 1 to 1331)
            Liu, Q.Y., Pandey, S., Singh, R.K., Lin, W., Ribocco, M.,
            Borow, Borowski, H., Smith, B., Leblanc, J., Walker, P.R. and
            Sikorska, M.
            DNasey: a rat DNaseI-like gene coding for a constitutively
            expressed chromatin-bound endonuclease
            Biochemistry 37 (28): 10134-10143 (1998)
            98332537
            PUBMED 9665719
REFERENCE 2 (bases 1 to 1331)
            Liu, Q.Y., Singh, R.K., Lin, W. and Sikorska, M.
            Direct Submission
            Submitted (24-DEC-1997) Institute for Biological Sciences, National
            Research Council, 1200 Montreal Road, Bldg. M-54, Box 4, Ottawa, ON
            K1A 0R6, Canada
FEATURES
    source
        1..1331
            /organism="Rattus norvegicus"
            /strain="Sprague-Dawley"
            /db_xref="taxon:10116"
            /tissue type="liver"
            153..1085
            /function="apoptotic endonuclease"
            /note="similar to rat DNase I; chromatin-bound"
            /codon_start=1
            /product="DNasey"
            /protein_id="AAC28937.1"
            /db_xref="GI:3395772"
            /translation="MSLYPASPLASLLFILAHGALSRLCSFNVRSPGSKENH
            NMDITVKIIRCDLILMEIKDSNNRIPILMEKINRNRSGITNYVVISRLGNT
            YKQVAFLYKELVSKSYHHYDGDADVSRFPVWQSPHTANKDVIPLHTTPT
            TPETSVKEIDELVEYTDVVKHKAENFIEMGDENAGCSYVFKKMKXIRLRITDRFV
            VWLIGQEDTIVKSTNCAYDRIVLAGEIVSVVSKSVDFQKAYKLTEEDALDVSDHFP
            VEFKLOSRAFTNSKSVTLKTKTKRSKS"
BASE COUNT 351 a 333 c 320 g 327 t
ORIGIN

```

Query Match	74.7%;	Score 501.4;	DB 10;	Length 1331;	
Best Local Similarity	85.3%;	Pred. No. 1.1e-125;			
Matches 559;	Conservative	0;	Mismatches 96;	Indels 0;	Gaps 0;
QY 1	ACAACAGGATCTGCCCATCTGATGGAGAGCTAAACGGAAATTCAGAAAGGCATAA	60			
Db	358 ACAACACATCTGCCCAGTCTGATGGAGAGCTAAACGGAAATTCAGAAAGGCATAA	417			
QY 61	CATACAACTATGTGATTAGTCTCGCTTGGAAAGAACACATATAAGAACAGTATGCCT	120			
Db	418 CATACAACTATGTGATTAGTCTCGCTTGGAAAGAACACATATAAGAACAGTATGCCT	477			
QY 121	TTCTCTATAAGAAAGCTAGTGTCTGTAACAAAGCTTACCTCTACCAAGCTATCAGG	180			
Db	478 TCCTCTACAGGAGAGCTGGTGTCTGTGAAGCAAAATACCTTACCATGACTATCAGG	537			
QY 181	CTGGAGACCGAGATGTTTTCAGGAGAACCTTGTGTGCTGTGTTCCAGTACCCCTACA	240			
Db	538 ATGGAGACAGACGCTGTTTTCAGGAGACCTTGTGTGTTTTCAGGCGCCCTTCA	597			
QY 241	CGCTCTCAAGGACTTCGTGATTGTGCTCCCTGCACACACCTCTGAGACATCCGTTAGAG	300			
Db	598 CTGCTCCCAAGGACTTCGTGATTGTGCTCCCTGCACACACCTCTGAAACCTCCGTTAAG	657			
QY 301	AGATTGATGAGCTGGTGTGATCTACACAGATGTGAACGCTCGCTGGAATTCAGAGAAAT	360			
Db	658 AGATAGATGAGCTGGTGTGATCTACACGATGTGAGAGACGATGGAAGGCAGAGAAAT	717			
QY 361	TCATTTTTCATGGTACTTCAATGCTGGCTGAGCTAGTCCCAAGAGGCTTGAAGG	420			
Db	718 TCATTTTTCATGGTACTTCAATGCTGGCTGAGCTAGTCCCAAGAGGCTTGAAGA	777			
QY 421	ACATCGCTGAGGACGACCCCAAGTTCGTTGGCTGATCGGGGACCAAGAGGACACCA	480			
Db	778 ACATCGCTTGGAGACAGACCCCAAGTTCGTTGGCTGATCGGGGACCAAGAGGACACCA	837			
QY 481	CGGTCAAGAGAGCAAACTGCGCTATGACAGGATCGTGTCTAGAGGACAAATATTG	540			
Db	838 CGGTCAAGAGAGCAACGCTGTGCTTATGACAGGATGTGCTTCCGCGACCAAGAGATAG	897			
QY 541	TCAACTCTGGTGTCTCAATCAAACTCGTCTTGAATTCAGAAAGCTTACAGTTGT	600			
Db	898 TCAACTCTGGTGTCTCAATCAAACTCGTCTTGAATTCAGAAAGCTTACAGTTGT	957			
QY 601	CTGAATCGAAGGCTCGGATGTGAGGACCACTTTCAGTTCATCATCATCA	655			
Db	958 CTGAAGAGAGGCTCGGATGTGAGGACCACTTTCAGTTCATCATCATCA	1012			
RESULT 10					
LOCUS	RNU75689	1417 bp	mRNA	linear	ROD 15-JUN-1998
DEFINITION	Rattus norvegicus DNase gamma mRNA, complete cds.				
ACCESSION	U75689				
VERSION	U75689.1	GI:3220024			
KEYWORDS	Rattus norvegicus.				
SOURCE	Rattus norvegicus				
ORGANISM	Rattus norvegicus				
REFERENCE	1 (bases 1 to 1417)				
AUTHORS	Shiokawa,D., Ohyama,H., Yamada,T., Takahashi,K. and Tanuma,S.				
TITLE	Identification of an endonuclease responsible for apoptosis in rat thymocytes				
JOURNAL	Eur. J. Biochem. 226 (1), 23-30 (1994)				
MEDLINE	95045594				
PUBMED	7957253				
REFERENCE	2 (bases 1 to 1417)				
AUTHORS	Shiokawa,D. and Tanuma,S.				
TITLE	Molecular cloning and expression of a cDNA encoding an apoptotic endonuclease DNase gamma				
Biochem. J. 332 (Pt 3), 713-720 (1998)					
98285539					
PUBMED	9620874				
REFERENCE	3 (bases 1 to 1417)				
AUTHORS	Shiokawa,D. and Tanuma,S.				
TITLE	Direct Submission				
JOURNAL	Submitted (22-OCT-1996) Biochemistry, Science Univ. of Tokyo,				
FEATURES	Shinjuku-Ku ichigaya funagawaracho, Tokyo 162, Japan				
Location/Qualifiers					
1..1417					
organism="Rattus norvegicus"					
/strain="Sprague-Dawley"					
/db_xref="taxon:10116"					
/tissue type="spleen"					
203..1135					
/function="apoptosis"					
/note="deoxyribonuclease"					
/codon_start=1					
/product="DNase gamma"					
/protein_id="AAC40134.1"					
/db_xref="GI:3220025"					
/translation="MSLVPSAPYLASLLLFILALHGLSLRLCSFNVSFGESKKENH					
NAMDIIVKIIKEDLILLMEIKDSNNICPMLEKNGNSRRSTTYNYVSRIGRNT					
YKEQAFDYKEKLVSRAKLYHDYQDGDVDFREPFVMPQAPFTAAKDFVIVPLH					
TPBTSVKEIDELADVTVRRRWKAENFIEMGFNAGCCSYVPKKAWKRIRLTPNF					
WMLGDQEDTTVKKSTSCAYDRIVLRGOEIVNSVVRSSGVDFQKAYELSEALDV					
SDHPVBFKLOSRAPFTNSRKSVSLKKKKGSRs"					
BASE COUNT	386 a 353 c 333 g 345 t				
ORIGIN					
Query Match	74.7%;	Score 501.4;	DB 10;	Length 1417;	
Best Local Similarity	85.3%;	Pred. No. 1.1e-125;			
Matches 559;	Conservative	0;	Mismatches 96;	Indels 0;	Gaps 0;
QY 1	ACAACAGGATCTGCCCATCTGATGGAGAGCTAAACGGAAATTCAGAAAGGCATAA	60			
Db	408 ACAACACATCTGTCCCATCTGATGGAGAGCTAAACGGAAATTCAGAAAGGCATAA	467			
QY 61	CATACAACTATGTGATTAGTCTCGCTTGGAAAGAACACATATAAGAACAGTATGCCT	120			
Db	468 CATACAACTATGTGATTAGTCTCGCTTGGAAAGAACACATATAAGAACAGTATGCCT	527			
QY 121	TTCTCTATAAGAAAGCTAGTGTCTGTGTAACAAAGCTTACCTTACCAAGCTATCAGG	180			
Db	528 TCCTCTACAGGAGAGCTGGTGTCTGTGAAGGCAAAATACCTTACCATGACTATCAGG	587			
QY 181	CTGGAGACCGAGATGTGTTTCCAGGAGACCTTGTGGTCTGGTTCAGTACCCCTACA	240			
Db	588 ATGGAGACACAGACGCTGTTTCCAGGAGACCTTGTGGTTCAGGAGCCCTTCA	647			
QY 241	CGCTCTCAAGGACTTCGTGATTGTGCTCCCTGCACACACCTCTGAGACATCCGTTAGAG	300			
Db	648 CTGCTGCCAAGGACTTCGTGATTGTGCTCCCTTGCACACAACTCTGAAACCTCCGTTAAG	707			
QY 301	AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAACGCTCGCTGGAATTCAGAGAAAT	360			
Db	708 AGATAGATGAGCTGGCTGACGCTTACACGAGTGTGAGAGACGATGGAAGGCAGAAAT	767			
QY 361	TCATTTTTCATGGTGAATTCATGCTGGCTGAGCTAGTCTCCCAAGAGGCTGGAAG	420			
Db	768 TCATTTTTCATGGTGAATTCATGCTGGCTGAGCTAGTCTCCCAAGAGGCTGGAAGA	827			
QY 421	ACATCGCTGAGGACGACCCCAAGTTCGTTGGCTGATCGGGGACCAAGAGGACACCA	480			
Db	828 ACATCGCTTGGAGACAGACCCCAAGTTCGTTGGCTGATCGGGGACCAAGAGGACACCA	887			
QY 481	CGGTCAAGAGAGACCAAACTCGCTTATGACAGGATCGTGTCTAGAGGACAAATATTG	540			
Db	888 CGGTCAAGAGAGACCAAGCTGTGCTATGACAGGATGTGCTTCCGAGACAGATAG	947			
QY 541	TCAACTCTGGTCTCTCAATCAAACTCGTCTTGAATTCAGAAAGCTTACAGTTGT	600			
Db	948 TCAACTCTGTGTTCCCGCTCCAGTGGCGTCTTTGACTTTTCAGAAAGCTTATGATTGT	1007			

```
QY 601 CTGAATCGAAGGCGCTGATGTCACGACCACTTTCAGTTCATCATCATCA 655
Db 1008 CTGAAGAGAGGCGCTGATGTCAGTGCACACTTTCAGTTCAGTTCAGTCA 1062

RESULT 11
LOCUS AR047845 1208 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 2 from patent US 5821103.
ACCESSION AR047845
VERSION AR047845.1 GI:5970188
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1208)
AUTHORS Tanuma,S.-i.
TITLE Deoxyribonuclease
JOURNAL Patent: US 5821103-A 2 13-OCT-1998;
FEATURES Location/Qualifiers
source 1..1208
BASE COUNT 319 a 304 c 283 g 302 t
ORIGIN
Query Match 71.7%; Score 481.2; DB 6; Length 1208;
Best Local Similarity 85.1%; Pred. No. 3.5e-120;
Matches 550; Conservative 0; Mismatches 93; Indels 3; Gaps 1;

QY 1 ACAACAGATCTGCCCTACTGATGAGAACTAAACGAAATTCAGAAAGCATTA 60
Db 217 ACAACAGATCTGCCCTACTGATGAGAACTAAACGAAATTCAGAAAGCATTA 276

QY 61 CATACAACTATGTGATGATCTCGCTTGGAGAAACATATTAAGAACATATGCT 120
Db 277 CATACAACTATGTGATGATCTCGCTTGGAGAAACATATTAAGAACATATGCT 336

QY 121 TTCTCTATAAGAAAGGCTAGTCTGTGTAACAAAGCTACCTTACGACATATCAG 180
Db 337 TTCTCTATAAGAAAGGCTAGTCTGTGTAACAAAGCTACCTTACGACATATCAG 396

QY 181 CTGAGAGCGAGATGTGTTTTCAGAGGAACCTTGTGCTCGTTCAGATCACCCTCA 240
Db 397 ATGAGAGCGAGATGTGTTTTCAGAGGAACCTTGTGCTCGTTCAGATCACCCTCA 456

QY 241 CCGCTGCAAGAGACTCTGTGATGTCCTCCCTGCAACACCCCTGAGACATCCGTTAG 300
Db 457 CTGCTGCCAAGAGACTCTGTGATGTCCTCCCTGCAACACCCCTGAGACATCCGTTAG 516

QY 301 AGATTGATGAGCTGCTGATGCTTACACAGATGTGAAACGTGCTGAATGACAGAA 360
Db 517 AGATTGATGAGCTGCTGATGCTTACACAGATGTGAAACGTGCTGAATGACAGAA 576

QY 517 AGATTGATGAGCTGCTGATGCTTACACAGATGTGAAACGTGCTGAATGACAGAA 576
Db 517 AGATTGATGAGCTGCTGATGCTTACACAGATGTGAAACGTGCTGAATGACAGAA 576

QY 361 TCATTTTCATGGGTGACTTCAATGTGGCTGACGTACGTCCCAAGAAAGCTTGAAG 420
Db 577 TCATTTTCATGGGTGACTTCAATGTGGCTGACGTACGTCCCAAGAAAGCTTGAAG 636

QY 421 ACATCCGCTGAGAGCGAGACCCCAAGTTCGTTTGGCTGATCGGGAGACCAAGAGACA 480
Db 637 ACATCCGCTGAGAGCGAGACCCCAAGTTCGTTTGGCTGATCGGGAGACCAAGAGACA 696

QY 481 CCGTCAAGAGAGCAAACTGCGCTTATGACAGATCGTGTCTAGAGACAAATATTTG 540
Db 697 CCGTCAAGAGAGCAAACTGCGCTTATGACAGATCGTGTCTAGAGACAAATATTTG 756

QY 541 TCAACTCTGGGTGCTCTCAATCAAACTGCTTGTGATTTCCAGAAACTTACAGTTGT 600
Db 757 TCAACTCTGGGTGCTCTCAATCAAACTGCTTGTGATTTCCAGAAACTTACAGTTGT 816

QY 601 CTGAATCGAAGGCGCTGATGTCAGTGCACACTTTCAGTTCATCATCA 643
Db 817 CTGAAGAGAGGCGCTGATGTCAGTGCACACTTTCAGTTCATCATCA 862
```

```
RESULT 12
LOCUS E11687 1208 bp RNA linear PAT 29-SEP-1997
DEFINITION cDNA encoding novel deoxyribonuclease (DNase) gamma which cut
specifically linker site of chromatin DNA.
ACCESSION E11687
VERSION E11687.1 GI:22025323
KEYWORDS JP 1996187079-A/1.
SOURCE Rattus rattus.
ORGANISM Rattus rattus.
REFERENCE 1 (bases 1 to 1208)
AUTHORS Tanuma,Y.
TITLE NEW DEOXYRIBONUCLEASE
JOURNAL Patent: JP 1996187079-A 1 23-JUL-1996;
TANUMA YASUKAZU
OS Rattus rattus (rat)
PN JP 1996187079-A/1
PD 23-JUL-1996
PF 06-SEP-1995 JP 1995255647
PR 06-SEP-1994 JP 94P 239518
PI TANUMA YASUKAZU
PC C12N9/16,C07K14/47,C07K16/40,C12N1/21,C12N15/09,(C12N1/21,PC
C12N1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH key location/Qualifiers
FT source 1..1208
FT /organism='Rattus rattus'
FT /tissue_type='Thymus, Spleen' FT
FT /cell_type='Cell nucleus'
FT 5'UTR 1..11
FT CDS 12..944
FT /product='Novel deoxyribonuclease (DNase) gamma
FT which cut
FT specifically linker site of chromatin DNA' FT
FT precursor RNA 12..86
FT 3'UTR 945..1208.
FEATURES Location/Qualifiers
source 1..1208
/organism='Rattus rattus'
/db_xref='taxon:10117'
BASE COUNT 319 a 304 c 283 g 302 t
ORIGIN
Query Match 71.7%; Score 481.2; DB 6; Length 1208;
Best Local Similarity 85.1%; Pred. No. 3.5e-120;
Matches 550; Conservative 0; Mismatches 93; Indels 3; Gaps 1;

QY 1 ACAACAGATCTGCCCTACTGATGAGAACTAAACGAAATTCAGAAAGCATTA 60
Db 217 ACAACAGATCTGCCCTACTGATGAGAACTAAACGAAATTCAGAAAGCATTA 276

QY 61 CATACAACTATGTGATGATCTCGCTTGGAGAAACATATTAAGAACATATGCT 120
Db 277 CATACAACTATGTGATGATCTCGCTTGGAGAAACATATTAAGAACATATGCT 336

QY 121 TTCTCTATAAGAAAGGCTAGTCTGTGTAACAAAGCTACCTTACGACATATCAG 180
Db 337 TTCTCTATAAGAAAGGCTAGTCTGTGTAACAAAGCTACCTTACGACATATCAG 396

QY 181 CTGAGAGCGAGATGTGTTTTCAGAGGAACCTTGTGCTCGTTCAGATCACCCTCA 240
Db 397 ATGAGAGCGAGATGTGTTTTCAGAGGAACCTTGTGCTCGTTCAGATCACCCTCA 456

QY 241 CCGCTGCAAGAGACTCTGTGATGTCCTCCCTGCAACACCCCTGAGACATCCGTTAG 300
```

```
Db 457 CTGCTGCAAGGACTTCGTGATGTCCTCCCTTGACACAACTCTCTGAAACCTCGTTAAAG 516
Qy 301 AGATTGATGAGCTGGGTGATGTGTACACAGATGTGAAACCGTCGCTGGAATGACAGAAATT 360
Db 517 AGATAGATGAGCTGGGTGACGTCTACACGATGTAGAGACCATGGAAGGCAGAGATT 576
Qy 361 TCAATTTTCATGGGTGATCTCAATGCTGGCTGACAGTACGTCCTCCCAAGAGGCCTGGAAGG 420
Db 577 TCAATCTTCATGGGTGATTTCAATGCTGGCTGACAGTACGTCCTCCCAAGAGGCCTGGAAGA 636
Qy 421 ACATCCGCTGAGGACGACCCCAAGTTCGTTGGCTGATCGGGGACCAAGAGACACCA 480
Db 637 ACATCCGTTTGAGGACAGACCCCAACTTTGTTGGCTGATGAGGACCAAGAGACACCA 696
Qy 481 CGGTCAAGAGAGACCAAACTGCGCTTATGACAGGATCGTCTTAGAGGACAAATATTG 540
Db 697 CGGTCAAGAGAGACCAAGCTGTGCTTATGACAGGATGTGCTTCGCGACAGAGATAG 756
Qy 541 TCAACTCTGTGCTCTCAATCAAACTCGTCTTTGATTTCCAGAAAGCTTACAGTTGT 600
Db 757 TCAACTCTGTGTTCCCGCTCCAGTGGCTCTTTGACTTTCAAGAAAGCTTATGAGTTGT 816
Qy 601 CTGAATCGAAGGCCCTG--GATGTCAGCAGCACCTTCCAGTTCA 643
Db 817 CTGAAGAGGAGGCCCTGGATGATGTGATGACCACTTTCCAGTTGA 862

RESULT 13
MMU76110
LOCUS Mus musculus DNase gamma mRNA, complete cds.
DEFINITION
ACCESSION U76110
VERSION U76110.1 GI:4098207
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE Shikawa, D., Hatanaka, T. and Tanuma, S.
JOURNAL Direct Submission
Submitted (24-Oct-1996) Biochemistry, Science Univ. of Tokyo,
Shinjuku-ku ichigaya funagawaracho, Tokyo 162, Japan
FEATURES
Location/Qualifiers
source 1..941
/organism="Mus musculus"
/db_xref="taxon:10090"
CDS 7..939
/function="apoptosis"
/feature="deoxyribonuclease"
/codon_start=1
/product="DNase gamma"
/db_xref="GI:4098208"
/translation="MSLHPASPRLASLLFILALHDTLALRCSFNVSFGASKKENH
EAMDIIKIKRCDLILMEIKDSSNICPMLKNGNSRESTTNYVVISRLGRNT
YKEQAFYVKEKLVSKYKHVHYDQDGTDFVSPBPVWFPHSPTAVKDFVPLH
TPETSKEIDELVDVYIDVRSQWKTEPFIMGDFNAGCSYVPRKAWQNRIRLTPKF
VWLIGQEDTIVKKSCTADRVLCQEIIVNSVVRSSGVDFQKAYDLSEEDALD
SDHFVPEFKQSRAFTNNRKSIVLKRKKGNRS"
BASE COUNT 255 a 229 c 228 g 229 t
ORIGIN
Query Match 70.8%; Score 474.8; DB 10; Length 941;
Best Local Similarity 82.9%; Pred. No. 1.9e-118;
Matches 542; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
Qy 2 CACACGATCTGCCCATCTGATGAGAGCTTAACGGAATTCAGAAAGGCATAC 61
Db 213 CAACAAATCTGTCCCATCTGATGAGAGCTGAATGAAATTCAGAAAGGCACAC 272
```

```
Qy 62 ATACAACATATGATAGTCTCGCTTGGAAAGAACACATATATAAGAACAGTATGCCTT 121
Db 273 ATACAACATATGATAGTCTCGCTTGGAAAGAACACATATATAAGAACAGTATGCCTT 332
Qy 122 TCTCTATAAAGAAAGCTAGTGTCTGTAAACAAAGCTACCTCTACACAGCTATCAGGC 181
Db 333 CGTCTACAAGGAGAGCTGTGTCTGTGAAGACAAATACCACCTACCATGCTATCAGGA 392
Qy 182 TGAGACGCGAGATGTGTTTCCAGGGAACCTTTGTGCTCCAGTTCACCTTACAC 241
Db 393 TGAGACACAGACGTTGTTTCCAGGAGGCCCTTTGTGTTTGGTTCCATTCCTTTTAC 452
Qy 242 CGTGTCAAGGACTTCGTGATTTGCCCTTGACACACCCCTGAGACATCCGTTAGAGA 301
Db 453 TGTGTCAAGGACTTCGTGATTTGCCCTTGACACAACTCCCGAGACCTCCGTTAAGA 512
Qy 302 GATTGATGAGCTGGCTGATGTCTACACAGATGTGAACCGTCTGCTGGAATGCAGAGAA 361
Db 513 GATGATGAGCTGGCTGATGTCTACCGATGTGAGAACCCAGTGGAGACAGAGAA 572
Qy 362 CATTTTCATGGGTGACTTCAATGCTGGCTGACGCTACGTCCTCCCAAGAGGCCTTGAAGA 421
Db 573 CATTTTCATGGGTGATTTCAACGCGCTGTAGTATGTCCTCCCAAGAGGCCTTGAAGA 632
Qy 422 CATTCGCTGAGGACGAGACCCCAAGTTCGTTGGCTGATCGGGGACCAAGAGACACAC 481
Db 633 CATTCGTTTGAAGACGAGACCCCAAGTTCGTTGGCTGATTTGGGGACCAAGAGACACTAC 692
Qy 482 GGTCAAGAGAGACAAACTGCGCTATCACAGGATCGCTTAGAGGACAAATATTGT 541
Db 693 GGTCAAGAGAGATACAGCTGTGCTTATCACAGGATGTGCTTTGTTGGACAGAGATAGT 752
Qy 542 CAACTCTGCTGCTCTCAATCAAACTCTCTTGTGATTTCCAGAAAGCTTACAGGTTGTC 601
Db 753 CAACTCTGCTGCTTCCCGCTTCCAGTGGCTCTTGTGACTTTCCAGAAAGCTTATGACTTGC 812
Qy 602 TGAATCAAGAGGCCTGGATGTCAGGACCACTTCCAGTTCATCATCATCA 655
Db 813 TGAGGAGGAGGCCTGGATGTCAGTGCATGATCCTTCCAGTTGAGTTTAAAGCTACA 866

RESULT 14
AF047355
LOCUS Mus musculus liver and spleen DNase precursor (LSD) mRNA, complete
DEFINITION cds.
ACCESSION AF047355
VERSION AF047355.1 GI:2905787
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 1124)
Cloning and characterization of an actin-resistant DNase I-like
endonuclease secreted by macrophages
JOURNAL Gene 215 (2), 291-301 (1998)
MEDLINE 98382522
PUBMED 9714828
REFERENCE 2 (bases 1 to 1124)
AUTHORS Baron, W.F., Pan, C.Q., Spencer, S.A., Ryan, A.M., Lazarus, R.A. and
Baker, K.P.
TITLE Direct Submission
JOURNAL Submitted (09-FEB-1998) Molecular Biology, Genentech Inc., 1 DNA
Way, South San Francisco, CA 94080, USA
FEATURES
Location/Qualifiers
source 1..1124
/organism="Mus musculus"
/db_xref="taxon:10090"
/gene="LSD"
```

```
CDS
173..1105
/ gene="LSD"
/ note="actin resistant activity; macrophage specific DNase
I; LS-DNase"
/ codon_start=1
/ product="liver and spleen DNase precursor"
/ protein_id="AAC35753.1"
/ db_xref="GI:2905788"
/ translation="MSLHPASPRLASLILFLALHDTLALRLCSFNVRSGASKENH
EAMDILVKIIRKCDLILMEIKDSSNNICPMLEKLNKNSRSTTYNVISSRLGRNT
YKEQYAFVYKEKLVSVKTKYHYDODDTPFSEPPVWFHSPFTAVKDFVILPLH
TTPETSVKEIDELVDVYDVSQWKTEFIEMGDFNAGCSYVKKAMQNIILRTDPKF
VMLIGDQEDTVKSTSCAYDRIVLCGEIIVNSVVPSSGVFDFQKAYDISEBALDV
SDHPFEPKLOSSRAFTNNRKSVLKKKKKNRS"
sig_peptide
173..247
/ gene="LSD"
/ evidence=not_experimental
248..1102
/ gene="LSD"
/ product="liver and spleen DNase"
mat_peptide
BASE COUNT 299 a 273 c 273 g 279 t
ORIGIN
Query Match 70.8%; Score 474.8; DB 10; Length 1124;
Best Local Similarity 82.9%; Pred. No. 1.9e-118;
Matches 542; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
```

```
BC012671
LOCUS BC012671 2110 bp mRNA linear ROD 07-AUG-2002
DEFINITION Mus musculus, similar to deoxyribonuclease 1-like 3, clone
MGC:13854 IMAGE:4160709, mRNA, complete cds.
ACCESSION BC012671
VERSION BC012671.1 GI:15215118
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
1 (bases 1 to 2110)
Strausberg, R.
AUTHORS Direct Submission
TITLE Submitted (15-AUG-2001) National Institutes of Health, Mammalian
JOURNAL Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT Contact: MGC help desk
Email: cgabs-rc@mail.nih.gov
DNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadansystemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
FEATURES
source
1..2110
/ organism="Mus musculus"
/ db_xref="taxon:10090"
/ map="FVB/N"
/ clone="MGC:13854 IMAGE:4160709"
/ tissue_type="Liver, normal. 5 month old male mouse."
/ clone_lib="NCI CGAP_L19"
/ lab_host="DH10B"
/ note="vector: pCMV-SPORT6"
151..1083
/ codon_start=1
/ product="Similar to deoxyribonuclease 1-like 3"
/ protein_id="AAH12671.1"
/ db_xref="GI:15215119"
/ translation="MSLHPASPRLASLILFLALHDTLALRLCSFNVRSGASKENH
EAMDILVKIIRKCDLILMEIKDSSNNICPMLEKLNKNSRSTTYNVISSRLGRNT
YKEQYAFVYKEKLVSVKTKYHYDODDTPFSEPPVWFHSPFTAVKDFVILPLH
TTPETSVKEIDELVDVYDVSQWKTEFIEMGDFNAGCSYVKKAMQNIILRTDPKF
VMLIGDQEDTVKSTSCAYDRIVLCGEIIVNSVVPSSGVFDFQKAYDISEBALDV
SDHPFEPKLOSSRAFTNNRKSVLKKKKKNRS"
BASE COUNT 526 a 581 c 464 g 539 t
ORIGIN
Query Match 70.5%; Score 473.2; DB 10; Length 2110;
Best Local Similarity 82.7%; Pred. No. 5.5e-118;
Matches 541; Conservative 0; Mismatches 113; Indels 0; Gaps 0;
```

Db 477 CGTCTACAAGGAGAACTGGTGTCTGTGAAGACAAATACCACTACCATGACTATCAGGA 536  
Qy 182 TGGAGACGAGATGTGTTTCCAGGAACCTTTGTGGTCTGGTTCAGTCAACCTTACAC 241  
Db 537 TGGAGACACAGACGCTGTTTCCAGGAGCCCTTTGTGGTTGGTTCCATTCCCCCTTAC 596  
Qy 242 CGCTGTCAAGGACTTCGTGATGTGTCCTCCCTGCACACCCCTGAGACATCCGTTAGAGA 301  
Db 597 TCGTGTCAAGGACTTCGTGATGTGTCCTCCCTTGACACAACCTCCCGAGACCTCCGTTAAGA 656  
Qy 302 GATTGATGAGCTGGCTGATGTCTACACAGATGTGAACGCTCGCTGGAATGCAGAGAAATT 361  
Db 657 GATAGATGAGCTGGTCTGATGTCTACACGATGTGAAGCCAGTGAAGACAGAGAAATT 716  
Qy 362 CATTTTCATGGGTGACTTCAATGCTGGCTGCAGCTACGTCCCAAGAGGCCCTGGAAGGA 421  
Db 717 CATCTTCATGGGTGATTTCAACGCGGCTGTAGCTATGTCCCCAAGAGGCCCTGGCAGAA 776  
Qy 422 CATCCGCTGAGGACGGACCCCAAGTTCTTTGGCTGATCGGGGACCAAGAGGACACCAC 481  
Db 777 CATTCGTTTGAGGACGGACCCCAAGTTCTTTGGCTGATTTGGGACCAAGAGGACACTAC 836  
Qy 482 GGTCAAGAAAGACACAACTGCGCCTATGACAGGATCGTGTAGAGGACAAATATTCT 541  
Db 837 GGTCAAGAAAGATACCAAGCTGTGCTTATGACAGGATTTGCTTTTGGGACAAAGATAGT 896  
Qy 542 CAACTCTGGTGTCTCAATCAAACTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGTC 601  
Db 897 CAACTCCGTGTTCCCGGTTCCAGTGGCGTCTTTGACTTTCAGAAAGCTTATCAGTTGTC 956  
Qy 602 TGAATCGAAGGCCCTGGATGTGACGACCACTTTCAGTTTCATCATCATCA 655  
Db 957 TGAGGAGGAGGCCCTGGATGTGAGTGCATCACTTCCAGTTGAGTTTAAGCTACA 1010

Search completed: December 9, 2002, 11:30:57  
Job time : 1421.41 secs

**This Page Blank (uspto)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 08:59:14 ; Search time 193.384 Seconds  
(without alignments)  
7813.939 Million cell updates/sec

Title: US-09-905-114-3

Perfect score: 671

Sequence: 1 acacagatctgcccata.....atcatcatgaagaaccatga 671

Scoring table: IDENTITY\_NUC

Gapop 10\_0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

1: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SID82/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SID82/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SID82/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SID82/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	671	100.0	671	24	AAD29089
2	592	88.2	592	24	AAD29088
3	539	80.3	578	24	AAS17900
4	527	78.5	1023	24	ABN96895
5	527	78.5	1023	24	ABL61914
6	527	78.5	1023	24	ABL64458
7	527	78.5	1023	24	ABL64961
8	527	78.5	1079	18	AAT74076
9	525.4	78.3	1108	24	ABK84429

10	518.4	77.3	556	24	AAS17901	Partial cDNA encod
11	479.6	71.5	1208	17	AAT17733	Deoxyribonuclease,
12	474.8	70.8	1124	18	AAT74083	Murine LS-DNase nu
13	348.4	51.9	350	24	AAS17902	partial cDNA encod
14	314.8	46.9	350	24	AAS17903	partial cDNA encod
15	300.6	44.8	349	24	AAS17904	Partial cDNA encod
16	195.6	29.2	392	24	ABN94132	Gene #630 used to
17	195.6	29.2	392	24	ABL62246	Colon adenocarcino
18	163.2	24.3	1039	15	AAQ54429	Human DNase I. HO
19	163.2	24.3	1039	22	ABA02661	Human DNase I.1.
20	161.6	24.1	783	22	ABA02662	Human DNase I.2.
21	161.6	24.1	858	22	ABA02663	Human DNase I.3.
22	161.6	24.1	1000	21	AAA28433	DNA encoding StrII
23	161.6	24.1	1039	11	AAQ05512	Sequence encoding
24	161.6	24.1	1039	15	AAQ05512	Human DNase gene.
25	161.6	24.1	1548	22	ABA02702	Humanised HMFG-1 F
26	161.6	24.1	1548	22	ABA02703	Humanised HMFG-1 F
27	161.6	24.1	1548	22	ABA02705	Humanised HMFG-1 F
28	161.6	24.1	1554	22	ABA02682	Humanised HMFG-1 F
29	161.6	24.1	1554	22	ABA02683	Humanised HMFG-1 F
30	161.6	24.1	1554	22	ABA02685	Humanised HMFG-1 F
31	161.6	24.1	1557	22	ABA02704	Humanised HMFG-1 F
32	161.6	24.1	1560	22	ABA02710	Humanised HMFG-1 F
33	161.6	24.1	1560	22	ABA02711	Humanised HMFG-1 F
34	161.6	24.1	1560	22	ABA02713	Humanised HMFG-1 F
35	161.6	24.1	1560	22	ABA02714	Humanised HMFG-1 F
36	161.6	24.1	1560	22	ABA02715	Humanised HMFG-1 F
37	161.6	24.1	1560	22	ABA02716	Humanised HMFG-1 F
38	161.6	24.1	1563	22	ABA02684	Humanised HMFG-1 F
39	161.6	24.1	1566	22	ABA02706	Humanised HMFG-1 F
40	161.6	24.1	1566	22	ABA02707	Humanised HMFG-1 F
41	161.6	24.1	1566	22	ABA02709	Humanised HMFG-1 F
42	161.6	24.1	1569	22	ABA02712	Humanised HMFG-1 F
43	161.6	24.1	1575	22	ABA02708	Humanised HMFG-1 F
44	161.6	24.1	1578	22	ABA02717	Humanised HMFG-1 F
45	161.6	24.1	1578	22	ABA02718	Humanised HMFG-1 F

#### ALIGNMENTS

RESULT 1  
AAD29089  
ID AAD29089 standard; DNA; 671 BP.  
XX  
AC AAD29089;  
XX  
DT 16-MAY-2002 (first entry)  
XX  
DE Bovine 22kDa recombinant FAA (rFAA) coding sequence.  
XX  
DE Bovine; fertility associated antigen; FAA; sperm fertility; acrosome;  
KW artificial insemination; plasma membrane; sperm cell; anti-infertility;  
KW reproductive tract; ds.  
XX  
OS Bos sp.  
XX  
PN WO200206444-A2.  
XX  
PD 24-JAN-2002.  
XX  
PF 16-JUL-2001; 2001WO-US22097.  
XX  
PR 14-JUL-2000; 2000US-218140P.  
XX  
PA (ARIZ-) ARIZONA BOARD OF REGENTS.  
XX  
PI Zhang H, Ax RL, Bellin ME;  
XX  
DR WPT; 2002-171803/22.  
XX  
PT Novel amino acid and polynucleotide sequence for fertility associated antigen useful for increasing stability of plasma membrane, acrosome



PT and other portions of sperm cell, and increasing fertility of mammals  
 XX  
 PS Claim 1; Fig 5; 37pp; English.  
 XX

CC The invention relates to bovine fertility associated antigen (FAA) and  
 CC its nucleic acid sequence. FAA is useful for increasing the stability of  
 CC the plasma membrane plus acrosome of a sperm cell and/or other portions  
 CC of a sperm cell, in particular mammalian sperm cell such as a sperm cell  
 CC from buffalo, cow, horse, mice, pig, sheep, human, avian sperm cell  
 CC including turkey or chicken sperm cell. FAA is also useful for increasing  
 CC the fertility of sperm from a mammal including buffalo, cow, horse, mice,  
 CC pig, sheep or human, or avian (chicken or turkey). FAA is added to the  
 CC suspension of sperm cells prior to preservation or prior to  
 CC administration of the sperm cells to a mammal in artificial insemination.  
 CC Alternatively, fertility of sperm from a mammal is increased by placing  
 CC FAA into a female's reproductive tract prior to deposition of the sperm  
 CC cells into the female's reproductive tract by copulation or artificial  
 CC insemination. The fertility of the male mammal is increased by injecting  
 CC FAA into the male mammal's reproductive tract. The present sequence is  
 CC bovine 22KDa recombinant FAA (rFAA) coding sequence.  
 CC  
 SO Sequence 671 BP; 195 A; 166 C; 154 G; 156 T; 0 other;

Query Match 100.0%; Score 671; DB 24; Length 671;  
 Best Local Similarity 100.0%; Pred. No. 4e-202;  
 Matches 671; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAACAGAGTCTGCCCTACTGATGAGAGCTTAACGGAATTCAGAAAAAGCATTA 60  
 DB 1 ACAACAGATCTGCCCTACTGATGAGAGCTTAACGGAATTCAGAAAAAGCATTA 60  
 QY 61 CATACACTATGTGATGATGCTCTGCTTGAAGAAACATATTAAGAAGATGCTCT 120  
 DB 61 CATACACTATGTGATGATGCTCTGCTTGAAGAAACATATTAAGAAGATGCTCT 120  
 QY 121 TTCTTATAAAGAAAGCTAGTGTGTTAAAAAGCTACTCTCAACCACTTTCAGG 180  
 DB 121 TTCTTATAAAGAAAGCTAGTGTGTTAAAAAGCTACTCTCAACCACTTTCAGG 180  
 QY 181 CTGAGACGCGAGATGTGTTTCCAGGAAACCTTGTGTGTTCCAGTACCCCTACA 240  
 DB 181 CTGAGACGCGAGATGTGTTTCCAGGAAACCTTGTGTGTTCCAGTACCCCTACA 240  
 QY 241 CCGCTGTCAAGACTTGTGATGTGCTCCCTGACACCAACCTTGAAGATCCGTTAAG 300  
 DB 241 CCGCTGTCAAGACTTGTGATGTGCTCCCTGACACCAACCTTGAAGATCCGTTAAG 300  
 QY 301 AGATTGATGACCTGGCTGATGCTACAGATGTAAGACGTGCTGGAATGAGAAATT 360  
 DB 301 AGATTGATGACCTGGCTGATGCTACAGATGTAAGACGTGCTGGAATGAGAAATT 360  
 QY 361 TCATTTTCATGGTGTACTTCAATGTGCTGAGCTACGCTCCCAAGAGGCTTGAAG 420  
 DB 361 TCATTTTCATGGTGTACTTCAATGTGCTGAGCTACGCTCCCAAGAGGCTTGAAG 420  
 QY 421 ACATCCGCTTGAAGACGAGACCCCAAGTTCGTTGCTGATGCGGAGCAACAGACCA 480  
 DB 421 ACATCCGCTTGAAGACGAGACCCCAAGTTCGTTGCTGATGCGGAGCAACAGACCA 480  
 QY 481 CGGTCAAAAGACCAAACTGCGCTATGACAGATGTGTTGAGAGCAAAAATATTG 540  
 DB 481 CGGTCAAAAGACCAAACTGCGCTATGACAGATGTGTTGAGAGCAAAAATATTG 540  
 QY 541 TCAACTCTGTGTGCTCTCAATCAAACTGCTTGTGATTTCCAGAAAGCTTACAGATTG 600  
 DB 541 TCAACTCTGTGTGCTCTCAATCAAACTGCTTGTGATTTCCAGAAAGCTTACAGATTG 600  
 QY 601 CTGAATCGAAGGCCCTGATGTCAAGCACTTCCAGTTTCATCATCATCATCATCATG 660  
 DB 601 CTGAATCGAAGGCCCTGATGTCAAGCACTTCCAGTTTCATCATCATCATCATCATG 660  
 QY 661 AAGAACCATGA 671

DB 661 AAGAACCATGA 671

# RESULT 2

ID AAD29088 standard; cDNA; 592 BP.

AC AAD29088;

DT 16-MAY-2002 (first entry)

DE Bovine fertility associated antigen (FAA) partial cDNA.

XX Bovine; fertility associated antigen; FAA; sperm fertility; acrosome;

KW artificial insemination; plasma membrane; sperm cell; anti-infertility;

KW reproductive tract; ss.

OS Bos sp.

XX Key Location/Qualifiers

FT CDS 1..591 /feature= a

FT /product= "Bovine fertility associated antigen (FAA)"

FT /note= "CDS does not include start and stop codon"

XX MO200206444-A2.

PD 24-JAN-2002.

PF 16-JUL-2001; 2001WO-US22097.

PR 14-JUL-2000; 2000US-218140P.

PA (ARIZ-) ARIZONA BOARD OF REGENTS.

XX Zhang H, Ax RL, Bellin ME;

XX WPI; 2002-171803/22.

DR P-PSDB; AAE18281.

PT Novel amino acid and polynucleotide sequence for fertility associated

PT antigen useful for increasing stability of plasma membrane, acrosome

PT and other portions of sperm cell, and increasing fertility of mammals

PS Claim 1; Fig 6; 37pp; English.

CC The invention relates to bovine fertility associated antigen (FAA) and

CC its nucleic acid sequence. FAA is useful for increasing the stability of

CC the plasma membrane plus acrosome of a sperm cell and/or other portions

CC of a sperm cell, in particular mammalian sperm cell such as a sperm cell

CC from buffalo, cow, horse, mice, pig, sheep, human, avian sperm cell

CC including turkey or chicken sperm cell. FAA is also useful for increasing

CC the fertility of sperm from a mammal including buffalo, cow, horse, mice,

CC pig, sheep or human, or avian (chicken or turkey). FAA is added to the

CC suspension of sperm cells prior to preservation or prior to

CC administration of the sperm cells to a mammal in artificial insemination.

CC Alternatively, fertility of sperm from a mammal is increased by placing

CC FAA into a female's reproductive tract prior to deposition of the sperm

CC cells into the female's reproductive tract by copulation or artificial

CC insemination. The fertility of the male mammal is increased by injecting

CC FAA into the male mammal's reproductive tract. The present sequence is

CC bovine fertility associated antigen (FAA) partial cDNA.

SO Sequence 592 BP; 170 A; 143 C; 142 G; 137 T; 0 other;

Query Match 88.2%; Score 592; DB 24; Length 592;

Best Local Similarity 100.0%; Pred. No. 4.1e-177;

Matches 592; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 GAGAAAGTAAAGGAAATTCAGAAAAAGCATTAACATCACTATGATGATGCTCTGCC 86

Db	1	GAGAAGCTAAACGGAAATTCAGAAAGGATACATACAACTATGTGATTAGCTCTCGC	60
QY	87	CTTGGGAAGAAACACATATAAGAACAGTATGCTTCTCTATAAGAAAAAGCTAGTGCT	146
Db	61	CTTGGGAAGAAACACATATAAGAACAGTATGCTTCTCTATAAGAAAAAGCTAGTGCT	120
QY	147	GTAATAACAAAGCTACCTCTACACGACTATCAGGCTGGAGACGAGATGTTTCCAGG	206
Db	121	GTAATAACAAAGCTACCTCTACACGACTATCAGGCTGGAGACGAGATGTTTCCAGG	180
QY	207	GAACCCCTTTGTGCTCTGCTTCCAGTACCCCTACACCGCTGTCAGGACTTCGTTGATGTC	266
Db	181	GAACCCCTTTGTGCTCTGCTTCCAGTACCCCTACACCGCTGTCAGGACTTCGTTGATGTC	240
QY	267	CCCTCTGCACACACCCCTCGATCGATCCGTTAGAGAGATTGATGAGCTGGCTGATGCTAC	326
Db	241	CCCTCTGCACACACCCCTCGATCGATCCGTTAGAGAGATTGATGAGCTGGCTGATGCTAC	300
QY	327	ACAGATGTGAACGCTCGCTGGAATGCAGAGAAATTTCAATTTTCATGGGTGACTTCAATGCT	386
Db	301	ACAGATGTGAACGCTCGCTGGAATGCAGAGAAATTTCAATTTTCATGGGTGACTTCAATGCT	360
QY	387	GGCTGAGCTACGCTCCCAAGAGGCTGGAAGGACATCCGCTGAGGACGACCCCAAG	446
Db	361	GGCTGAGCTACGCTCCCAAGAGGCTGGAAGGACATCCGCTGAGGACGACCCCAAG	420
QY	447	TTGCTTTGGCTGATCGGGACCAAGAGGACACACCGTCAAGAGAGACAAACTCGGC	506
Db	421	TTGCTTTGGCTGATCGGGACCAAGAGGACACACCGTCAAGAGAGACAAACTCGGC	480
QY	507	TATGACAGGATCGTGTAGAGGACAAAATATTTGTCAACTCTGCTGCTCAATCAAAAC	566
Db	481	TATGACAGGATCGTGTAGAGGACAAAATATTTGTCAACTCTGCTGCTCAATCAAAAC	540
QY	567	CTGCTCTTTGATTTCCAGAAAGCTTACAGTTGTTGTAATCGAAGGCCCTGG	618
Db	541	CTGCTCTTTGATTTCCAGAAAGCTTACAGTTGTTGTAATCGAAGGCCCTGG	592
RESULT 3			
AAS17900			
ID	AAS17900 standard; cDNA; 578 BP.		
XX	AC AAS17900;		
XX	AC AAS17900;		
DT	26-MAR-2002 (first entry)		
XX	DNA encoding bovine fertility-associated antigen (FAA).		
DE	Fertility-associated antigen; FAA; antiinfertility; sperm cell acrosome;		
KW	reproductive fitness; single nucleotide polymorphism; SNP; bovine; cow;		
KW	ss.		
XX	Bos taurus.		
OS			
XX	Key		
FT	Location/Qualifiers		
FT	1..578		
FT	/*tag= a		
FT	/product= "FAA"		
FT	/note= "Fertility-associated antigen"		
FT	/partial		
FT	/note= "No start or stop codon given"		
FT	/transl_except= (pos:289..300, aa:A)		
FT	/note= This codon has an apparent 9 nucleotide insertion		
FT	/transl_except= (pos:325..326, aa:R)		
FT	/note= "This codon has an apparent 1 nucleotide deletion which alters the reading frame"		
FT	/transl_except= (pos:576..578, aa:RLSBSKAL)		
XX			
PN	W0200176529-A2.		
XX			

PD	18-OCT-2001.		
XX			
PF	09-APR-2001; 2001WO-US10802.		
XX			
PR	07-APR-2000; 2000US-195225P.		
XX			
PA	(ZHAN/) ZHANG H M.		
PA	(AXRL/) AX R L.		
XX			
PI	Zhang HM, Ax RL;		
XX			
DR	WPI: 2002-066307/09.		
DR	P-PSDB; AAU11516.		
XX			
PT	Novel human or bovine fertility-associated antigen useful for		
PT	stabilising sperm cell acrosome and increasing fertility of a male		
PS	Claim 25; Fig 2A-E; 54pp; English.		
XX			
CC	The invention describes a novel isolated human fertility-associated		
CC	antigen (FAA) and a bovine FAA polypeptide. Human or bovine FAA is useful		
CC	for increasing the stability of a sperm cell acrosome and for increasing		
CC	the fertility of a human male by administering FAA. Detecting single		
CC	nucleotide polymorphisms (SNPs) of the human or bovine FAA gene is useful		
CC	for assaying fertility by detecting the presence or absence of a FAA SNP		
CC	in the mammal and correlating the presence or absence with the fertility		
CC	of the mammal. The FAA SNPs are chosen from FAA SNPs 1-34 given in the		
CC	specification. FAA SNP is also useful for determining the reproductive		
CC	fitness of a human or bovine, by determining the nucleotide sequence of		
CC	the FAA SNP, quantifying the fertility of more than one mammal containing		
CC	the FAA SNP and correlating the frequency of the FAA SNP to the		
CC	reproductive fitness of the mammals. FAA improves the integrity of sperm		
CC	membranes and increases the capacitation of sperm derived from either		
CC	fertile or infertile humans. Early genetic identification of infertility		
CC	improves the prognosis for subsequent attempts at fertilisation and		
CC	facilitates early intervention to determine whether the individual may		
CC	benefit from fertility treatments to avoid costly and/or emotional		
CC	problems with attempted inseminations. This sequence encodes the bovine		
CC	fertility-associated antigen (FAA) described in the method of the		
CC	invention.		
XX			
SQ	Sequence 578 BP; 167 A; 140 C; 138 G; 133 T; 0 other;		
Query Match 80.3%; Score 539; DB 24; Length 578;			
Best Local Similarity 98.3%; Pred. No. 2.6e-160;			
Matches 569; Conservative 0; Mismatches 0; Indels 10; Gaps 2;			
QY	27	GAGAAGCTAAACGGAAATTCAGAAAGGACATACATCAACTATGTGATTAGCTCTCGC	86
Db	1	GAGAAGCTAAACGGAAATTCAGAAAGGACATACATCAACTATGTGATTAGCTCTCGC	60
QY	87	CTTGGGAAGAAACACATATAAGAACAGTATGCTTCTCTATAAGAAAAAGCTAGTGCT	146
Db	61	CTTGGGAAGAAACACATATAAGAACAGTATGCTTCTCTATAAGAAAAAGCTAGTGCT	120
QY	147	GTAATAACAAAGCTACCTCTACACGACTATCAGGCTGGAGACGACAGATGTTTCCAGG	206
Db	121	GTAATAACAAAGCTACCTCTACACGACTATCAGGCTGGAGACGACAGATGTTTCCAGG	180
QY	207	GAACCCCTTTGTGCTCTGCTTCCAGTACCCCTACACCGCTGTCAGGACTTCGTTGATGTC	266
Db	181	GAACCCCTTTGTGCTCTGCTTCCAGTACCCCTACACCGCTGTCAGGACTTCGTTGATGTC	240
QY	267	CCCTCTGCACACACCCCTCGAGATCCGTTAGAGAGATTGATGAGCTGGCTGAGCTGGCT	317
Db	241	CCCTCTGCACACACCCCTCGAGATCCGTTAGAGAGATTGATGAGCTGGCTGAGCTGGCT	300
QY	318	GATGCTCTACACAGATGTGAAACGCTCGCTGGAATGCAGAGAAATTTCAATTTTCATGGGTGAC	377
Db	301	GATGCTCTACACAGATGTGAAACGCTC-CTGGAATGCAGAGAAATTTCAATTTTCATGGGTGAC	359
QY	378	TTCAATGCTGCTGAGCTACGTCCTCCCAAGAGGCTCGGAAGGACATCCGCTCGAGGACG	437

Db 360 TTCATGCTGCGTGCAGCTACGTCCTCCCAAGAGGCGCTGGAAGGACATCCGCTGAGAGG 419

Oy 438 GACCCCAAGTTGTTGGCTGATCGGGGACCAAGAGGACACCGGTCAAGAAAGCACA 497

Db 420 GACCCCAAGTTGTTGGCTGATCGGGGACCAAGAGGACACCGGTCAAGAAAGCACA 479

Oy 498 AACTGCGCTATGACGAGATCGTGCTTAAGAGCAAAATATTGTCAACTCTGCTGCT 557

Db 480 AACTGCGCTATGACGAGATCGTGCTTAAGAGCAAAATATTGTCAACTCTGCTGCT 539

Oy 558 CAATCAAACTGCTGTTGATTTCCAGAAAGCTTACAG 596

Db 540 CAATCAAACTGCTGTTGATTTCCAGAAAGCTTACAG 578

RESULT 4

ABN96895

ID ABN96895 standard; DNA; 1023 BP.

XX

AC ABN96895;

XX

DT 13-AUG-2002 (first entry)

XX

DE Gene #3393 used to diagnose liver cancer.

XX

KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;

KW metastatic liver tumour; cytostatic; expression profile; disease state;

KW disease progression; drug toxicity; drug efficacy; drug metabolism.

XX

OS Homo sapiens.

XX

PN WO200229103-A2.

XX

PD 11-APR-2002.

XX

PF 02-OCT-2001; 2001WO-US30589.

XX

PR 02-OCT-2000; 2000US-237054P.

XX

PA (GENE-) GENE LOGIC INC.

XX

PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;

XX

DR WPI; 2002-426119/45.

XX

PT Diagnosing and detecting the progression of liver cancer.

PT hepatocellular carcinoma or metastatic liver tumor in a patient,

PT involves detecting the level of expression of two or more genes in a

PT liver tissue sample

XX

PS Claim 1; SEQ ID NO 3393; 298bp; English.

XX

XX The invention relates to a novel method for diagnosing and detecting the

CC progression of liver cancer, hepatocellular carcinoma or metastatic liver

CC tumour in a patient, and differentiating metastatic liver cancer from

CC hepatocellular carcinoma in a patient, involving detecting the level of

CC expression of two or more genes represented in ABN93503-ABN97455 in a

CC tissue sample. The method of the invention has hepatotropic, and

CC cytostatic activity. The method is useful for diagnosing and detecting

CC the progression of liver cancer, hepatocellular carcinoma and metastatic

CC liver carcinoma in a patient. The method is useful for identifying

CC expression profiles which serve as useful diagnostic markers as well as

CC markers that can be used to monitor disease states, disease progression,

CC drug toxicity, drug efficacy and drug metabolism.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX

XX Sequence 1023 BP; 312 A; 244 C; 240 G; 227 T; 0 other;

Query Match 78.5%; Score 527; DB 24; Length 1023;

Best Local Similarity 87.8%; Pred. No. 2.2e-156;

Matches 575; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

Oy 1 ACAACGAGATTTGCCCATCTACTGATGAGAAAGCTPAAACGAAATTCAGAAAGGACATA 60

Db 215 ACAACGAGATTTGCCCATCTACTGATGAGAAAGCTPAAACGAAATTCAGAAAGGACATA 274

Oy 61 CATCAACTATGATGATTAAGCTCTGCTGGAAGAAACACATATTAAGAACATATGCTT 120

Db 275 GGTACAACTATGATGATTAAGCTCTGCTGGAAGAAACACATATTAAGAACATATGCTT 334

Oy 121 TTCTCTATTAAGAAAGGCTAGTGTCTGTAAACAAAGCTACCTTACACAGCATATGCTT 180

Db 335 TTCTCTATTAAGAAAGGCTAGTGTCTGTAAACAAAGCTACCTTACACAGCATATGCTT 394

Oy 181 CTGAGAGCAGATGATGTTTTCAGGGAACCTTTGTGCTGCTTCACTACCTTACA 240

Db 395 ATGAGAGCAGATGATGTTTTCAGGGAACCTTTGTGCTGCTTCACTACCTTACA 454

Oy 241 CGGCTGTCAAGAGCTTGTGATGTTGCTCCCTGCAACACCCCTGAGACATCCGTTAGG 300

Db 455 CTGCTGTCAAGAGCTTGTGATGTTATCCCTGCAACACCCCTGAGACATCCGTTAGG 514

Oy 301 AGATTGATGAGCTGCTGATGCTTACACAGATGTAACGTCGCTGGAATTCAGAAATT 360

Db 515 AGATTGATGAGCTGCTGATGCTTACACAGATGTAACGTCGCTGGAATTCAGAAATT 574

Oy 361 TCATTTTTCATGAGTGAATCTTCAATGCTGCTGACGCTACGTCCTCCAGAAAGGCTGAGG 420

Db 575 TCATTTTTCATGAGTGAATCTTCAATGCTGCTGACGCTACGTCCTCCAGAAAGGCTGAGG 634

Oy 421 ACATCCGCTGAGAGGAGGAGCCCAAGTTGCTTTGGCTGATCGGGGACCAAGAGGACCA 480

Db 635 ACATCCGCTGAGAGGAGGAGCCCAAGTTGCTTTGGCTGATCGGGGACCAAGAGGACCA 694

Oy 481 CGGTCAAGAAAGGAGCAAACTGCGCTTATGACAGATGCTGCTTAAAGGACAAATATTG 540

Db 695 CGGTCAAGAAAGGAGCAAACTGCGCTTATGACAGATGCTGCTTAAAGGACAAATATTG 754

Oy 541 TCACTCTGCTGCTGCTTCAATCAAACTCTGCTTTGATTTTTCAGAAAGCTTACAGTTGT 600

Db 755 TCACTCTGCTGCTGCTTCAATCAAACTCTGCTTTGATTTTTCAGAAAGCTTACAGTTGT 814

Oy 601 CTGAATGGAAGGCGCTGATGTCAGGACCACTTCAGTTCAATCATCATCA 655

Db 815 CTGAATGGAAGGCGCTGATGTCAGGACCACTTCAGTTCAATCATCATCA 869

RESULT 5

ABL61914

ID ABL61914 standard; DNA; 1023 BP.

XX

AC ABL61914;

XX

DT 15-MAY-2002 (first entry)

XX

DE Colon adenocarcinoma related gene sequence SEQ ID NO:251.

XX

KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;

KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;

KW cytostatic; gene therapy; antineoplastic; Wilms tumour; adenocarcinoma;

XX

OS Homo sapiens.

XX

PN WO200194629-A2.

XX

PD 13-DEC-2001.

XX

PF 30-MAY-2001; 2001WO-US10838.

XX

PR 05-JUN-2000; 2000US-209473P.

PR 05-JUN-2000; 2000US-209531P.

PR 18-SEP-2000; 2000US-233133P.

PR 18-SEP-2000; 2000US-233617P.

20-SEP-2000; 2000US-234009P.  
 20-SEP-2000; 2000US-234034P.  
 20-SEP-2000; 2000US-234052P.  
 22-SEP-2000; 2000US-234509P.  
 22-SEP-2000; 2000US-234567P.  
 25-SEP-2000; 2000US-234923P.  
 25-SEP-2000; 2000US-234924P.  
 25-SEP-2000; 2000US-235077P.  
 25-SEP-2000; 2000US-235082P.  
 25-SEP-2000; 2000US-235134P.  
 25-SEP-2000; 2000US-235280P.  
 26-SEP-2000; 2000US-235637P.  
 26-SEP-2000; 2000US-235638P.  
 27-SEP-2000; 2000US-235711P.  
 27-SEP-2000; 2000US-235720P.  
 27-SEP-2000; 2000US-235840P.  
 27-SEP-2000; 2000US-235863P.  
 28-SEP-2000; 2000US-236028P.  
 28-SEP-2000; 2000US-236032P.  
 28-SEP-2000; 2000US-236033P.  
 28-SEP-2000; 2000US-236034P.  
 28-SEP-2000; 2000US-236109P.  
 28-SEP-2000; 2000US-236111P.  
 29-SEP-2000; 2000US-236842P.  
 29-SEP-2000; 2000US-236891P.  
 02-OCT-2000; 2000US-237172P.  
 02-OCT-2000; 2000US-237173P.  
 02-OCT-2000; 2000US-237278P.  
 02-OCT-2000; 2000US-237294P.  
 02-OCT-2000; 2000US-237295P.  
 02-OCT-2000; 2000US-237316P.  
 03-OCT-2000; 2000US-237425P.  
 03-OCT-2000; 2000US-237598P.  
 03-OCT-2000; 2000US-237604P.  
 03-OCT-2000; 2000US-237606P.  
 03-OCT-2000; 2000US-237608P.  
 01-NOV-2000; 2000US-244867P.  
 01-NOV-2000; 2000US-245084P.  
 (AVAL-) AVALON PHARM.  
 Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
 Soppet DR, Weaver Z;  
 WPI; 2002-188264/24.  
 Screening for anti-neoplastic agent involves exposing cells to a  
 chemical agent to be tested for anti-neoplastic activity, and  
 determining a change in expression of a gene of a signature gene set -  
 Claim 1; SEQ ID 251; 44pp; English.  
 The present invention describes a method (M1) for screening for an  
 anti-neoplastic agent. The method involves exposing cells to a chemical  
 agent to be tested for anti-neoplastic activity, determining a change in  
 expression of at least one gene (I) of a signature gene set, where (I)  
 comprises a sequence (S) selected from 8447 sequences (given in ABL6164  
 to ABL70110), or is at least 95% identical to (S), where a change in  
 expression is indicative of anti-neoplastic activity. (I) has cytostatic  
 activity and can be used in gene therapy. M1 can be used for screening  
 an anti-neoplastic agent, and can be used for producing a product which  
 is the data collected with respect to the anti-neoplastic agent as a  
 result of M1, and the data is sufficient to convey the chemical  
 structure and/or properties of the agent. M1 can be used in the  
 treatment of cancer such as colon, breast, stomach, lung, thyroid,  
 oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
 adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
 infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
 carcinoma, papillary carcinoma and Wilm's tumour.  
 Sequence 1023 BP; 312 A; 244 C; 240 G; 227 T; 0 other;  
 Query Match 78.5%; Score 527; DB 24; Length 1023;

Best Local Similarity 87.8%; Pred. NO. 2.2e-156;  
 Matches 575; Conservative 0; Mismatches 80; Indels 0; Gaps 0;  
 QY 1 ACAACAGGATCTGCCCTACTGATGGAGAGCTAAACCGNAATTCAAGAAAGGCATAA 60  
 DB 215 ACAACAGGATCTGCCCTACTGATGGAGAGCTAAACCGNAATTCAAGAGAGGCATAA 274  
 QY 61 CATACAACTATGTGATTAGCTCTCGCCTTTGGAGAAACACATATAAAGAACAGTATGCCT 120  
 DB 275 CGTACAACCTATGTGATTAGCTCTCGCCTTTGGAGAAACACATATAAAGAACATATGCCT 334  
 QY 121 TTCTCTATAAAGAAAGCTAGTGTCTGTAAACAAGCTACCTCTACACGAGCTATCAGG 180  
 DB 335 TTCTCTACAAGGAAAGCTGGTGTCTGTGAAGAGGAGTTTATCACTACCATGACTATCAGG 394  
 QY 181 CTGGAGACGCAGATGTTTTCAGGGAACCTTTGTGTCTGTTTCCAGTCAACCTTAC 240  
 DB 395 ATGGAGACGCAGATGTTTTCAGGGAACCTTTGTGTCTGTTTCCAGTCAACCTTAC 454  
 QY 241 CGCCTGTCAAGGACTTCGTGATTGTCTCCCTGCACACACCCCTGAGACATCCGTTAGAG 300  
 DB 455 CTGCTGTCAAGACTTCGTGATTATCCCTCTGCACACACCCAGAGACATCCGTTAAGG 514  
 QY 301 AGATTGATGAGCTGCTGATGTCTACACAGATGTGAACGTCGCTGGAATGCAGAAATT 360  
 DB 515 AGATCGATGAGTTGGTTGAGGTCTACACGAGCTGAAAACACCGCTGGAAAGCGGAGATT 574  
 QY 361 TCATTTTCATGGTGACATTCATGCTGCTGAGTACGCTCCCAAGAGAGGCTGGAAGG 420  
 DB 575 TCATTTTCATGGTGACATTCATGCTGCTGAGTACGCTCCCAAGAGAGGCTGGAAGG 634  
 QY 421 ACATCCGCTGAGGACGACCCCAAGTTCTGTTGGCTGATCGGGACCAAGAGGACACCA 480  
 DB 635 ACATCCGCTGAGGACTGACCCAGGTTTGTGCTGATCGGGACCAAGAGGACACCA 694  
 QY 481 CGGTCAAGAGAGACAAACTCGGCTATGACAGGATCGTCTTAGAGGACAAATATTG 540  
 DB 695 CGGTGAAGAGAGACCAACTGTGCATATGACAGGATTGTGCTTAGAGGACAAATCG 754  
 QY 541 TCAACTCTGCTGCTCTCAATCAAACTCGTCTTTGATTTCCAGAAAGCTTACAGTTCT 600  
 DB 755 TCAGTTCTGTTTCTCCCAAGTCAAAAGTGTGTTGACTTCCAGAAAGCTTACAGCTGA 814  
 QY 601 CTGAATCGAAGGCCCTGGATGTCAGGACCACTTTCCAGTTTCATCATCATCA 655  
 DB 815 CTGAAGAGAGGCCCTGGATGTCAGGACCACTTTCCAGTTTCATCATCA 869  
 RESULT 6  
 ID ABL64458 standard; DNA; 1023 BP.  
 XX ABL64458;  
 AC ABL64458;  
 XX  
 DT 15-MAY-2002 (first entry)  
 XX  
 DE Stomach cancer related gene sequence SEQ ID NO:2795.  
 XX  
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
 KW gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200194629-A2.  
 XX  
 PD 13-DEC-2001.  
 XX  
 PF 30-MAY-2001; 2001WO-US10838.  
 XX  
 PR 05-JUN-2000; 2000US-209473P.  
 PR 05-JUN-2000; 2000US-209531P.

PR 18-SEP-2000; 2000US-233133P.  
 PR 18-SEP-2000; 2000US-233617P.  
 PR 20-SEP-2000; 2000US-234009P.  
 PR 20-SEP-2000; 2000US-234034P.  
 PR 20-SEP-2000; 2000US-234052P.  
 PR 22-SEP-2000; 2000US-234509P.  
 PR 22-SEP-2000; 2000US-234567P.  
 PR 22-SEP-2000; 2000US-234923P.  
 PR 25-SEP-2000; 2000US-234924P.  
 PR 25-SEP-2000; 2000US-235077P.  
 PR 25-SEP-2000; 2000US-235082P.  
 PR 25-SEP-2000; 2000US-235134P.  
 PR 25-SEP-2000; 2000US-235280P.  
 PR 26-SEP-2000; 2000US-235637P.  
 PR 26-SEP-2000; 2000US-235638P.  
 PR 27-SEP-2000; 2000US-235711P.  
 PR 27-SEP-2000; 2000US-235720P.  
 PR 27-SEP-2000; 2000US-235840P.  
 PR 27-SEP-2000; 2000US-235863P.  
 PR 28-SEP-2000; 2000US-236028P.  
 PR 28-SEP-2000; 2000US-236032P.  
 PR 28-SEP-2000; 2000US-236033P.  
 PR 28-SEP-2000; 2000US-236034P.  
 PR 28-SEP-2000; 2000US-236109P.  
 PR 28-SEP-2000; 2000US-236111P.  
 PR 29-SEP-2000; 2000US-236842P.  
 PR 29-SEP-2000; 2000US-236891P.  
 PR 02-OCT-2000; 2000US-237122P.  
 PR 02-OCT-2000; 2000US-237173P.  
 PR 02-OCT-2000; 2000US-237278P.  
 PR 02-OCT-2000; 2000US-237294P.  
 PR 02-OCT-2000; 2000US-237295P.  
 PR 02-OCT-2000; 2000US-237316P.  
 PR 03-OCT-2000; 2000US-237425P.  
 PR 03-OCT-2000; 2000US-237598P.  
 PR 03-OCT-2000; 2000US-237604P.  
 PR 03-OCT-2000; 2000US-237606P.  
 PR 03-OCT-2000; 2000US-237608P.  
 PR 01-NOV-2000; 2000US-244867P.  
 PR 01-NOV-2000; 2000US-245084P.  
 XX (AVALON PHARM.  
 PA Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
 PI Soppe DR, Weaver Z;  
 XX WPI; 2002-188264/24.  
 DR  
 XX  
 XX Screening for anti-neoplastic agent involves exposing cells to a  
 PT chemical agent to be tested for anti-neoplastic activity, and  
 PT determining a change in expression of a gene of a signature gene set -  
 XX  
 PS Claim 1; SEQ ID 2795; 44pp; English.  
 XX  
 CC The present invention describes a method (M1) for screening for an  
 CC anti-neoplastic agent. The method involves exposing cells to a chemical  
 CC agent to be tested for anti-neoplastic activity, determining a change in  
 CC expression of at least one gene (1) of a signature gene set, where (1)  
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
 CC to ABL70110), or is at least 95% identical to (S), where a change in  
 CC expression is indicative of anti-neoplastic activity. (1) has cytostatic  
 CC activity and can be used in gene therapy. M1 can be used for screening  
 CC an anti-neoplastic agent, and can be used for producing a product which  
 CC is the data collected with respect to the anti-neoplastic agent as a  
 CC structure and/or properties of the agent. M1 can be used in the  
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer.  
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
 CC carcinoma, papillary carcinoma and Wilm's tumour.  
 CC  
 XX Sequence 1023 BP; 312 A; 244 C; 240 G; 227 T; 0 other;

Query Match 78.5%; Score 527; DB 24; Length 1023;  
 Best Local Similarity 87.8%; Pred. No. 2,2e-156;  
 Matches 575; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 1 ACAACAGATCTGCCCATCTGATCTGATGAGAACTTAAACGGAATTTCAAGAAAGGCATTA 60  
 DB 215 ACAACAGATCTGCCCATCTGATCTGATGAGAACTTAAACGGAATTTCAAGAGGCATTA 274  
 QY 61 CATAAACTATGATGATAGCTCTGCTGGAAGAAACATATATAAGAACAGATATCCT 120  
 DB 275 CGTAACTATGATGATAGCTCTGCTGGAAGAAACATATATAAGAACAGATATCCT 334  
 QY 121 TTCTCTATAAGAAAGCTAGTGTCTGTAAACAAAGCTACCTTACACGACTATCAGG 180  
 DB 335 TTCTCTATAAGAAAGCTAGTGTCTGTAAAGAGAGATTAACATACATGACTATCAGG 394  
 QY 181 CTGAGAGCGAGATGTTTTCAGGGAACCTTTGAGTCTGTTCCAGTCACTTACCA 240  
 DB 395 ATGAGAGCGAGATGTTTTCAGGGAACCTTTGAGTCTGTTCCAGTCACTTACCA 454  
 QY 241 CCGCTGTCAAGGACTTGTGATTTGCCCTGCACACACCCCTGAGACATCCGTAGAG 300  
 DB 455 CTGCTGTCAAGGACTTGTGATTTGCCCTGCACACACCCCTGAGACATCCGTAGAG 514  
 QY 301 AGATTGATGAGCTGCTGATGCTTACACAGATGTGAAACCTGCTGAATGACAGAAATT 360  
 DB 515 AGATCGATGAGTGTGAGTGTACACAGACGCTGAAACACCGCTGGAAGCGGAGATT 574  
 QY 361 TCATTTTCATGGGAGCTTCAATGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGAG 420  
 DB 575 TCATTTTCATGGGAGCTTCAATGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGAG 634  
 QY 421 ACATCCGCTGAGAGCGAGACCCCAAGTTCGTTGGCTGATCGGGAGCAAGAGAGACCA 480  
 DB 635 ACATCCGCTGAGAGCGAGACCCCAAGTTCGTTGGCTGATCGGGAGCAAGAGAGACCA 694  
 QY 481 CGGTCAAGAGAGCAAACTGCGCTTATGACAGATGCTGTTAGAGCAAAATTTTG 540  
 DB 695 CGGTCAAGAGAGCAAACTGCGCTTATGACAGATGCTGTTAGAGCAAAATTTTG 754  
 QY 541 TCAACTCGTGTGCTGCTCAATCAAACTGCTTGTGATTCGAGAAAGCTTACAGTGTG 600  
 DB 755 TCAAGTCTGTTGTTCCCAAGTCAACAGTGTGTTTGACTTCCAGAAAGCTTACAGTGA 814  
 QY 601 CTGAATCGAAGCCCTGATGTCAGGACCACTTTCAGTTCATCATCATCATCA 655  
 DB 815 CTGAAGAGAGAGCCCTGATGTCAGGACCACTTTCAGTTCATCATCATCATCA 869

RESULT 7  
 ABL64961  
 ID ABL64961 standard; DNA; 1023 BP.  
 XX  
 AC ABL64961;  
 XX  
 DT 15-MAY-2002 (first entry)  
 XX  
 DE Lung cancer related gene sequence SEQ ID NO:3298.  
 XX  
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; carcinos;  
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
 XX gene; ds.  
 OS Homo sapiens.  
 XX  
 PN M0200194629-A2.  
 XX  
 PD 13-DEC-2001.  
 XX  
 PF 30-MAY-2001; 2001MO-US10838.  
 XX

PR 05-JUN-2000; 2000US-209473P.  
PR 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 18-SEP-2000; 2000US-233617P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 20-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 22-SEP-2000; 2000US-234567P.  
PR 25-SEP-2000; 2000US-234923P.  
PR 25-SEP-2000; 2000US-234924P.  
PR 25-SEP-2000; 2000US-235077P.  
PR 25-SEP-2000; 2000US-235082P.  
PR 25-SEP-2000; 2000US-235134P.  
PR 25-SEP-2000; 2000US-235280P.  
PR 26-SEP-2000; 2000US-235637P.  
PR 26-SEP-2000; 2000US-235638P.  
PR 27-SEP-2000; 2000US-235711P.  
PR 27-SEP-2000; 2000US-235720P.  
PR 27-SEP-2000; 2000US-235840P.  
PR 27-SEP-2000; 2000US-235863P.  
PR 28-SEP-2000; 2000US-236028P.  
PR 28-SEP-2000; 2000US-236032P.  
PR 28-SEP-2000; 2000US-236033P.  
PR 28-SEP-2000; 2000US-236034P.  
PR 28-SEP-2000; 2000US-236109P.  
PR 28-SEP-2000; 2000US-236111P.  
PR 29-SEP-2000; 2000US-236842P.  
PR 29-SEP-2000; 2000US-236891P.  
PR 02-OCT-2000; 2000US-237172P.  
PR 02-OCT-2000; 2000US-237173P.  
PR 02-OCT-2000; 2000US-237278P.  
PR 02-OCT-2000; 2000US-237294P.  
PR 02-OCT-2000; 2000US-237295P.  
PR 02-OCT-2000; 2000US-237316P.  
PR 03-OCT-2000; 2000US-237425P.  
PR 03-OCT-2000; 2000US-237598P.  
PR 03-OCT-2000; 2000US-237604P.  
PR 03-OCT-2000; 2000US-237606P.  
PR 03-OCT-2000; 2000US-237608P.  
PR 01-NOV-2000; 2000US-244867P.  
PR 01-NOV-2000; 2000US-245084P.  
XX (AVAL-) AVALON PHARM.  
PA Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
XX Soppet DR, Weaver Z;  
PI WPI; 2002-188264/24.  
XX  
PT Screening for anti-neoplastic agent involves exposing cells to a  
PT chemical agent to be tested for anti-neoplastic activity, and  
PT determining a change in expression of a gene of a signature gene set  
XX  
PS Claim 1; SEQ ID 3298; 44pp; English.  
XX  
CC The present invention describes a method (M1) for screening for an  
CC anti-neoplastic agent. The method involves exposing cells to a chemical  
CC agent to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 847 sequences (given in ABL61664  
CC to ABL70110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
CC activity and can be used in gene therapy. M1 can be used for screening  
CC an anti-neoplastic agent, and can be used for producing a product which  
CC is the data collected with respect to the anti-neoplastic agent as a  
CC result of M1, and the data is sufficient to convey the chemical  
CC structure and/or properties of the agent. M1 can be used in the  
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
CC carcinoma, papillary carcinoma and Wilm's tumour.

XX SQ Sequence 1023 BP; 312 A; 244 C; 240 G; 227 T; 0 other;  
Query Match 78.5%; Score 527; DB 24; Length 1023;  
Best Local Similarity 87.8%; Pred. No. 2.2e-156;  
Matches 575; Conservative 0; Mismatches 80; Indels 0; Gaps 0;  
QY 1 ACAACAGGATCTGCCCATCTACTGATGGAGAGCTAAACGAAATTCAGAAAGGCAATAA 60  
DB 215 ACAACAGGATCTGCCCATCTACTGATGGAGAGCTGAACAGAAATTCAGAGAGGCAATAA 274  
QY 61 CATACAACTATGTAGTAGCTCTCCCTTGGAGAGAACACATATAAAGAACAGTAGTCCT 120  
DB 275 CGTCAAACTATGTAGTAGCTCTCCCTTGGAGAGAACACATATAAAGAACAGTAGTCCT 334  
QY 121 TTCTCTATAAAGAAAGCTAGTGTCTGTAAACAAAGCTACCTCTACCAAGCTATCATCAG 180  
DB 335 TTCTCTACAGGAAAGCTGGTGTCTGTGAGAGGAGTTATCACTACCATGACTATCAGG 394  
QY 181 CTGAGAGCGCAGATGTGTTTCCAGGGAACCCCTTTGTGTGTTCCAGTCAACCCCTACA 240  
DB 395 ATGGAGACGACAGTGTGTTTCCAGGAGGCCCTTTGTGTGTTTCCAATCTCCCCACA 454  
QY 241 CCGCTGTCAAGGACTTGTGATGTTGCCCTCTGCAACACCCCTCGAGACATCCGTTAGAG 300  
DB 455 CTGCTGTCAAAGACTTCGTGATTATCCCTCTGCACACCCACCCAGAGACATCCGTTAAGG 514  
QY 301 AGATTGATGAGCTGGCTGATGCTACACAGATGTGAACGTCGCTGGAATGCAGAGAATT 360  
DB 515 AGATCGATGAGTTGGTTGAGGTCATACCGAGCTGAACACACCGCTGGAGGCGGAGAA 574  
QY 361 TCATTTTTCATGGTGACTTCAATGCTGCTGAGCTAGCTCCCAAGAGGCTCTGGAAGG 420  
DB 575 TCATTTTTCATGGTGACTTCAATGCTGCTGAGCTAGCTCCCAAGAGGCTCTGGAAGG 634  
QY 421 ACATCCGCTGAGGACGGAACCCAAAGTTGCTTTGCTGATCGGGGACCAAGAGGACACCA 480  
DB 635 ACATCCGCTTGAGGACTGACCCAGGTTGTTTGGCTGATCGGGGACCAAGAGGACACCA 694  
QY 481 CGGTCAAGAAGACACAACTCGCGCTATGACAGGATCGTCTTAGAGGACAAATATTTG 540  
DB 695 CGTGAAGAAGAGCACCAACTGTGCTATGACAGGATGCTGCTTAGAGGACAAAGAAATCG 754  
QY 541 TCAACTCTGGTGGTCTCTCAATCAAACTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600  
DB 755 TCAGTTCTGTGTTCCCAAGTCAACACTGTTTGTGACTTCCAGAAAGCTTACAGGCTGA 814  
QY 601 CTGAATCGAAGGCCCTGGATGTACGACACCACTTTTCCAGTTTCATCATCATCA 655  
DB 815 CTGAAGAGGAGGCCCTGGATGTACGACGACCACTTTCCAGTTGAATTTAAACTACA 869  
RESULT 8  
AAT74076  
ID AAT74076 standard; cDNA; 1079 BP.  
XX AC AAT74076;  
XX DT 29-JAN-1998 (first entry)  
XX DE Human LS-DNase nucleotide sequence.  
XX KW DNase: actin; DNA viscoelasticity; systemic lupus erythematosus;  
KW cystic fibrosis; meningitis; pulmonary disease; gene therapy;  
KW recombinant preparation; DNA-hydrolytic activity; ss.  
XX OS Homo sapiens.  
XX PH Key Location/Qualifiers  
FT CDS 71..988  
FT sig\_peptide /\*tag= a  
FT 71..130  
FT /\*tag= b

mat\_peptide 131..985  
/\*tag= C  
/product= LS-DNase

PN W09728266-A1.  
XX 07-AUG-1997.  
PD 07-AUG-1997.  
XX 03-FEB-1997; 97WO-US01506.  
PF 03-FEB-1997; 97WO-US01506.  
XX 05-FEB-1996; 96US-0597078.  
PR 05-FEB-1996; 96US-0597078.  
XX (GETH ) GENENTECH INC.  
PA Baker KP, Baron WF;  
PI Baker KP, Baron WF;  
XX WPI: 1997-402619/37.  
DR P-PSDB; AAM23542.  
XX WPI: 1997-402619/37.  
DR P-PSDB; AAM23542.  
XX New isolated LS-DNase which is resistant to inhibition by actin -  
PT used for the treatment of e.g systemic lupus erythematosus and  
PT pulmonary diseases and disorders such as cystic fibrosis  
XX  
PS Claim 1; Fig 1; 32pp; English.

CC This DNA sequence encodes a novel human LS-DNase (AAM23542), which has  
CC DNA-hydrolytic activity but is resistant to inhibition by actin. It can  
CC be used for reducing the viscoelasticity or viscous consistency of  
CC DNA-containing material, as well as treatment of systemic lupus  
CC erythematosus or pulmonary disease of disorders such as cystic fibrosis,  
CC bronchitis, pneumonia, bronchiectasis, emphysema, asthma, tuberculosis  
CC or fungal infections. It can also be used for adjunctive treatment of  
CC abscesses or severe closed-space infections in conditions such as  
CC empyema, meningitis, peritonitis, sinusitis, otitis, periodontitis,  
CC pericarditis, pancreatitis, cholangitis, endocarditis and  
CC septic arthritis, as well as in topical treatments in a variety of  
CC inflammatory and infected lesions, e.g. of the skin and mucosal  
CC membranes, surgical wounds, ulcerative lesions and burns. LS-DNase may  
CC improve the efficacy of antibiotics used in the treatment of such  
CC infections (e.g. gentamicin activity is markedly reduced by reversible  
CC binding to intact DNA). It can also be used for the treatment of other  
CC non-infected conditions in which there is an accumulation of cellular  
CC debris that includes cellular DNA, such as pyelonephritis and  
CC tubulo-interstitial kidney disease. It can also be used in vitro  
CC diagnostic assays of a viscous material (e.g. sputum) from a patient.  
CC The anti-LS-DNase antibodies can be used for detection and purification  
CC of LS-DNase.  
XX  
XX Sequence 1079 BP; 330 A; 260 C; 252 G; 237 T; 0 other;

Query Match 78.5%; Score 527; DB 18; Length 1079;  
Best Local Similarity 87.8%; Pred. No. 2.3e-156;  
Matches 575; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 1 ACAACGAGATCTGCCCATCTGATGAGAGCTAAACGAAATTCAGAAAAGCATTA 60  
DB 261 ACAACGAGATCTGCCCATCTGATGAGAGCTAAACGAAATTCAGAAAAGCATTA 320  
QY 61 CATACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
DB 321 CATACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 380  
QY 121 TTCTCTATTAAGAAAGCTAGTGTCTGTAACAAAGCTACTTACACGACTATCAG 180  
DB 381 TTCTCTATTAAGAAAGCTAGTGTCTGTAACAAAGCTACTTACACGACTATCAG 440  
QY 181 CTGAGAGCCGACATGTTTCCAGGAAACCTTTGTGTGTGTTCCAGTACCTTACA 240  
DB 441 ATGGAGAGCCGACATGTTTCCAGGAAACCTTTGTGTGTGTTCCAGTACCTTACA 500  
QY 241 CCGGCTCAAGAGCTTCGATGTTCCCTTCACACACCCCTGAGACATCGTTAAG 300  
DB 501 CTGCTGTCAAAAGCTTCGATGTTTCCCTTCACACACCCCTGAGACATCGTTAAG 560

QY 301 AGATTGATGAGCTGGCTGATGCTTACACAGATGTAACAGTGGAAAGCTGGTAATGACAGAAATT 360  
DB 561 AGATTGATGAGCTGGCTGATGCTTACACAGATGTAACAGTGGAAAGCTGGTAATGACAGAAATT 620  
QY 361 TCATTTTCATGAGGAGCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
DB 621 TCATTTTCATGAGGAGCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 680  
QY 421 ACATCCGCTGAGGAGCGAGCCCAAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
DB 681 ACATCCGCTGAGGAGCTGAGCCCAAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 740  
QY 481 CGGTCAAGAGAGCAAACTGCGCTTATGACAGATGCTGCTTATGAGGACAAATATTTG 540  
DB 741 CGGTCAAGAGAGCAAACTGCGCTTATGACAGATGCTGCTTATGAGGACAAATATTTG 800  
QY 541 TCAACTCTGCT 600  
DB 801 TCAACTCTGCT 860  
QY 601 CTGAATCGAAGGCGCTGATGTCAGCAGCACCTTCCAGTTTCATCATCATCATCA 655  
DB 861 CTGAATCGAAGGCGCTGATGTCAGCAGCACCTTCCAGTTTCATCATCATCATCA 915

RESULT 9  
ABK84429  
ID ABK84429 standard; cDNA; 1108 BP.  
XX ABK84429;  
AC ABK84429;  
XX  
DT 14-AUG-2002 (first entry)  
XX  
DE Human cDNA differentially expressed in granulocytic cells #1000.  
XX  
KW Human; se; granulocytic cell; DNA chip; bacterial infection;  
KW viral infection; parasitic infection; protozoal infection;  
KW fungal infection; sterile inflammatory disease; psoriasis;  
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;  
KW cardiac reperfusion injury; renal reperfusion injury; AIDS;  
KW adult respiratory distress syndrome; inflammatory bowel disease;  
KW Crohn's disease; ulcerative colitis; peridomonal disease;  
KW granulocyte activation; chronic inflammation; allergy.  
OS Homo sapiens.  
XX  
XX W0200228999-A2.  
XX  
XX 11-APR-2002.  
XX  
XX 03-OCT-2001; 2001WO-US30821.  
XX  
XX 03-OCT-2000; 2000US-237189P.  
XX  
XX (GENE-) GENE LOGIC INC.  
XX  
XX Beazer-Barclay Y, Weisman SM, Yamaga S, Vockley J;  
PI WPI: 2002-435328/46.  
XX  
XX WPI: 2002-435328/46.  
XX  
XX Detecting granulocyte activation by detecting differential expression  
XX of genes associated with granulocyte activation, which serves as  
XX diagnostic markers that is useful for monitoring disease states and  
XX drug toxicity -  
XX  
XX Claim 1; SEQ ID No 1000; 114pp; English.  
XX  
XX The invention relates to detecting (M1) granulocyte (GC) activation  
XX (GCA), by detecting the level of expression of gene(s) (Gs) identified by  
XX DNA chip analysis as given in the specification, and comparing  
XX the expression level to an expression level in an unactivated  
XX GC, where differential expression of Gs is indicative of GCA.



CC Also included are modulating (M2) CA by contacting GC with an agent  
CC that alters the expression of at least one gene in Gs; (2) screening (M3)  
CC for an agent capable of modulating GCA or an inflammation (especially  
CC chronic) in a tissue, an allergic response in a subject, exposure of a  
CC subject to a pathogen or sterile inflammatory disease using the  
CC gene expression profile; (3) detecting (M4) an inflammation (especially  
CC chronic) in a tissue, an allergic response in a subject, exposure of a  
CC subject to a pathogen or sterile inflammatory disease, by detecting the  
CC level of expression in a sample of the tissue of gene(s) from Gs, where  
CC the level of expression of the gene is indicative of inflammation;  
CC (4) treating (M5) an inflammation (especially chronic) in a tissue,  
CC an allergic response in a subject, exposure of a subject to a pathogen  
CC or sterile inflammatory disease, by contacting a tissue having  
CC inflammation with an agent that modulates the expression of gene(s)  
CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for  
CC modulating GCA; M3 is useful for screening an agent capable of modulating  
CC GCA preferably in an inflammation in a tissue; M4 is useful for  
CC detecting an inflammation (especially chronic) in a tissue, an allergic  
CC response in a subject, exposure of a subject to a pathogen or sterile  
CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,  
CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal  
CC reperfusion injury, ARDS, adult respiratory distress syndrome,  
CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,  
CC periodontal disease; also bacterial infection, viral infection, and  
CC parasitic infection, protozoal infection, fungal infection and M5 is  
CC useful for treating one of the above conditions. The present  
CC sequence represents a gene differentially expressed in granulocytes.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX  
SQ Sequence 1108 BP; 329 A; 261 C; 271 G; 247 T; 0 other;

Query Match 78.3%; Score 525.4; DB 24; Length 1108;  
Best Local Similarity 87.6%; Pred. No. 7.5e-156;  
Matches 574; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 1 ACAACAGGATCTGCCCATCTGATGAGAAGCTAAACGGAAATTCAGAAAGGCATAA 60  
DB 283 ACAACAGGATCTGCCCATCTGATGAGAAGCTAAACGGAAATTCAGAGAGGCATAA 342  
QY 61 CATACAACTATGTGATTAGCTCTCGCCTTGGGAAGAAACACATATAAAGACAGTATGCCT 120  
DB 343 CGTACAACTATGTGATTAGCTCTCGCCTTGGGAAGAAACACATATAAAGACATATGCCT 402  
QY 121 TTCTCTATAAGAAAGCTAGTGTCTGTAAGAAACAAAGCTACTTACCACGACTATCAGG 180  
DB 403 TTCTCTATAAGAAAGCTAGTGTCTGTAAGAAAGAGGAGTATCACTACCATGACTATCAGG 462  
QY 181 CTGGAGACGCAGATGTGTTTCCAGGAAACCCCTTTGTGCTGTTCCAGTCAACCCCTACA 240  
DB 463 ATGGAGACGCAGATGTGTTTCCAGGAGCCCTTTGTGCTGTTCCATCTCCCCACA 522  
QY 241 CCGCTCTAAGGACTTGTGATGTGTCCTCCCTGCACACCCCTCGAGACATCGTGTAGAG 300  
DB 523 CTGCTCTAAGGACTTGTGATGTGTCCTCCCTGCACACCCCTCGAGACATCGTGTAGAG 582  
QY 301 AGATTGATGAGCTGCTGATCTACACAGATGTGAACGTCGCTGATGACAGAAAT 360  
DB 583 AGATGATGAGTGTGTTGAGTGTCTACACGAGCGTGAACACCGCTGGAAGCGGCGAAAT 642  
QY 361 TCATTTTCATGGTGACTTCAATGTGCTGAGTACGTGCTCCCAAGAGCGCTGGAAGG 420  
DB 643 TCATTTTCATGGTGACTTCAATGTGCTGAGTACGTGCTCCCAAGAGCGCTGGAAGG 702  
QY 421 ACATCCGCTGAGGACGACCCCAAGTTCCTTTGGCTGATCGGGGACCAAGAGCACCA 480  
DB 703 ACATCCGCTGAGGACTGACCCCAAGTTCCTTTGGCTGATCGGGGACCAAGAGCACCA 762  
QY 481 CGGTCAAGAGACCAAACTGCGCTTATGACAGGATCGTGTCTAGAGACAAATATTG 540  
DB 763 CGGTGAAGAAGACCAAACTGTGCTATGACAGGATTTGTGCTTAGAGGACCAAGAAATCG 822

QY 541 TCAACTCTGGTGGTCTCTCAATCAAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGGTTGT 600  
DB 823 TCAGTTCTCTGTTCTCCCAAGTCAAAACAGTGTGTTTGTGACTTCCAGAAAGCTTACAGCTGA 882  
QY 601 CTGAATCGAAGGCCCTGGATGTCAGGACCACTTCCAGTTCATCATCATCATCA 655  
DB 883 CTGAAGAGAGAGGCCCTGGATGTCAGGACCACTTCCAGTTCATCATCATCA 937  
RESULT 10  
AAS17901  
ID AAS17901 standard; cDNA; 556 BP.  
XX  
AC AAS17901;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Partail cDNA encoding human fertility-associated antigen (FAA), HCl.  
KW Fertility-associated antigen; FAA; antiinfertility; sperm cell acrosome;  
KW reproductive fitness; single nucleotide polymorphism; SNP; human;  
KW HCl; ss.  
XX Homo sapiens.  
OS  
FH Key Location/Qualifiers  
CDS 1..555  
FT /\*tag= a  
FT /product= "Human FAA"  
FT /note= "Human fertility-associated antigen, HCl"  
FT /partial  
FT /note= "No start or stop codon given"  
FT /transl\_except= (pos:1..3, aa:KVIRKCDIILVWEIKDS)  
FT /transl\_except= (pos:253..255, aa:S)  
XX WO200176529-A2.  
XX  
PD 18-OCT-2001.  
XX  
PF 09-APR-2001; 2001WO-US10802.  
XX  
PR 07-APR-2000; 2000US-195225P.  
XX  
XX (ZHAN/) ZHANG H M.  
PA (AXEL/) AX R L.  
XX  
PI Zhang HM, Ax RL;  
XX  
DR WPI; 2002-066307/09.  
DR P-PSDB; AAU11517.  
XX  
PT Novel human or bovine fertility-associated antigen useful for  
PT stabilising sperm cell acrosome and increasing fertility of a male  
XX  
PS Claim 1; Fig 2A-E; 54pp; English.  
XX  
CC The invention describes a novel isolated human fertility-associated  
CC antigen (FAA) and a bovine FAA polypeptide. Human or bovine FAA is useful  
CC for increasing the stability of a sperm cell acrosome and for increasing  
CC the fertility of a human male by administering FAA. Detecting single  
CC nucleotide polymorphisms (SNPs) of the human or bovine FAA gene is useful  
CC for assaying fertility by detecting the presence or absence of a FAA SNP  
CC in the mammal and correlating the presence or absence with the fertility  
CC of the mammal. The FAA SNPs are chosen from FAA SNPs 1-34 given in the  
CC specification. FAA SNP is also useful for determining the reproductive  
CC fitness of a human or bovine, by determining the nucleotide sequence of  
CC the FAA SNP, quantifying the fertility of more than one mammal containing  
CC the FAA SNP and correlating the frequency of the FAA SNP to the  
CC reproductive fitness of the mammals. FAA improves the integrity of sperm  
CC membranes and increases the capacitation of sperm derived from either  
CC fertile or infertile humans. Early genetic identification of infertility  
CC improves the prognosis for subsequent attempts at fertilisation and



CC facilitates early intervention to determine whether the individual may  
CC benefit from fertility treatments to avoid costly and/or emotional  
CC problems with attempted inseminations. This sequence encodes the human  
CC fertility-associated antigen (FAA), HCl, one of 4 partial prostate CDNA  
CC clones of the novel FAA gene described in the method of the invention.  
XX

SO Sequence 556 BP; 159 A; 130 C; 142 G; 125 T; 0 other;

Query Match 77.3%; Score 518.4; DB 24; Length 556;  
Best Local Similarity 96.2%; Pred. No. 8.5e-154;  
Matches 531; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
OY 1 ACAAGAGATGTCGCCCTACTGATGAGAGAGCTAAAGGAAATTCAGAAAGGATTA 60
D 5 ACAAGAGATGTCGCCCTACTGATGAGAGAGCTGAAACGAAATTCAGAGAGGATTA 64
OY 61 CATACAACTATGTGATTAGCTCTCCCTTGAAGAAACATATATAAGACATATGCTT 120
D 65 COTACAACTATGTGATTAGCTCTCCCTTGAAGAAACATATATAAGACATATGCTT 124
OY 121 TTCTCTATAAGAAAGCTAGTGTCTGTAACAAAGCTACTTACACGACTATCAGG 180
D 125 TTCTCTATAAGAAAGCTAGTGTCTGTAACAAAGCTACTTACACGACTATCAGG 184
OY 181 CTGAGAGCGCAATGTGTTTCCAGGAAACCTTTGTGTCTGTCTCCAGTACCCCTACA 240
D 185 ATGGAGACCGCAATGTGTTTCCAGGAAACCTTTGTGTCTGTCTCCAGTACCCCTACA 244
OY 241 CCGCTGTCAAGACTTGTGATTGTCCTCCCTGACACACCAACCCCTGAGACATCCGTTAG 300
D 245 CCGCTGTCAAGACTTGTGATTGTCCTCCCTGACACACCAACCCCTGAGACATCCGTTAG 304
OY 301 AGATTGATGAGCTGGCTGATGCTTACACAGATGTAAGAGTGGCTGGAATGACAGATT 360
D 305 AGATTGATGAGCTGGCTGATGCTTACACAGATGTAAGAGTGGCTGGAATGACAGATT 364
OY 361 TCATTTTATGGGTGACTTCATGCTGCGGTGAGCTAGCTCCCAAGAGGCTGGAAG 420
D 365 TCATTTTATGGGTGACTTCATGCTGCGGTGAGCTAGCTCCCAAGAGGCTGGAAG 424
OY 421 ACATCCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
D 425 ACATCCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 484
OY 481 CCGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
D 485 CCGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 544
OY 541 TCAACTCTGTTG 552
D 545 TCAACTCTGTTG 556
```

RESULT 11

AA117733  
ID AA117733 standard; cDNA, 1208 BP.

AA117733;

17-OCT-1996 (first entry)

Deoxyribonuclease, DNase-gamma cDNA.

XX Deoxyribonuclease; DNase-gamma; selective cleavage; linker;  
XX chomatin DNA; cell nucleus; 3'-hydroxy; 5'-phosphate; apoptosis;  
XX carcinogenesis; AIDS development; diagnosis; prevention;  
XX treatment; cancer; autoimmune disease; viral infections; ds.

OS Homo sapiens.

XX Key Location/Qualifiers  
FH 12..944  
FT CDS /\*tag= a

FT /note= "GAT codon comprising bases 837-839 has  
FT no corresponding amino acid in AAR94021"

XX MO9607735-A1.

XX 14-MAR-1996.

XX 06-SEP-1995; 95WO-JP01775.

XX 06-SEP-1994; 94JP-0239518.

XX (TANU/) TANUMA S.

XX Tanuma S;

XX WPI; 1996-171610/17.

XX P-PSDB; AAR94021.

PT DNase which selectively cuts linker region of chromatin DNA -  
PT useful in diagnosis, treatment and prevention of cancer, autoimmune  
PT diseases, viral infections, etc.

PS Claim 4; Pages 55-57; 75pp; Japanese.

XX The present sequence encodes the deoxyribonuclease, DNase-gamma,  
XX which is able to selectively cleave the linker part of chromatin  
XX DNA. DNase-gamma has a mol. wt. of 32000 by SDS-PAGE, or 30000 by  
XX gel filtration, an optimum pH of 5.6, is found in the cell nucleus,  
XX is not dependent on divalent cations, has a Zn(2+) inhibition  
XX IC50) > 1mM and leaves 3'-OH and 5'-P termini after cleavage. The  
XX DNase can be used at a mol. level to resolve the control  
XX mechanisms of carcinogenesis, the autoimmune system and AIDS  
XX treatment of cancer, autoimmune disease, viral infections, etc.,  
XX and as an apoptosis control agent.

SO Sequence 1208 BP; 319 A; 304 C; 282 G; 303 T; 0 other;

Query Match 71.5%; Score 479.6; DB 17; Length 1208;  
Best Local Similarity 85.0%; Pred. No. 2.6e-141;  
Matches 549; Conservative 0; Mismatches 94; Indels 3; Gaps 1;

```
OY 1 ACAAGAGATGTCGCCCTACTGATGAGAGAGCTAAAGGAAATTCAGAAAGGATTA 60
D 217 ACAAGAGATGTCGCCCTACTGATGAGAGAGCTAAAGGAAATTCAGAAAGGATTA 276
OY 61 CATACAACTATGTGATTAGCTCTCCCTTGAAGAAACATATATAAGACATATGCTT 120
D 277 CATACAACTATGTGATTAGCTCTCCCTTGAAGAAACATATATAAGACATATGCTT 336
OY 121 TTCTCTATAAGAAAGCTAGTGTCTGTAACAAAGCTACTTACACGACTATCAGG 180
D 337 TTCTCTATAAGAAAGCTAGTGTCTGTAACAAAGCTACTTACACGACTATCAGG 396
OY 181 CTGAGAGCGCAATGTGTTTCCAGGAAACCTTTGTGTCTGTCTCCAGTACCCCTACA 240
D 397 ATGGAGACCGCAATGTGTTTCCAGGAAACCTTTGTGTCTGTCTCCAGTACCCCTACA 456
OY 241 CCGCTGTCAAGACTTGTGATTGTCCTCCCTGACACCAACCCCTGAGACATCCGTTAG 300
D 457 CCGCTGTCAAGACTTGTGATTGTCCTCCCTGACACCAACCCCTGAGACATCCGTTAG 516
OY 301 AGATTGATGAGCTGGCTGATGCTTACACAGATGTAAGAGTGGCTGGAATGACAGATT 360
D 517 AGATTGATGAGCTGGCTGATGCTTACACAGATGTAAGAGTGGCTGGAATGACAGATT 576
OY 361 TCATTTTATGGGTGACTTCATGCTGCGGTGAGCTAGCTCCCAAGAGGCTGGAAG 420
D 577 TCATTTTATGGGTGACTTCATGCTGCGGTGAGCTAGCTCCCAAGAGGCTGGAAG 636
OY 421 ACATCCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
D 637 ACATCCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 696
```



```

FH Key Location/Qualifiers
FT CDS 1..349
FT /*tag= a
FT /product= "Human FFA"
FT /note= "Human fertility-associated antigen, HC2"
FT /partial
FT /note= "No start or stop codon given"
XX
XX MO200176529-A2.
XX
XX 18-OCT-2001.
XX
XX 09-APR-2001; 2001WO-US10802.
XX
XX 07-APR-2000; 2000US-195225P.
XX
XX (ZHAN/) ZHANG H M.
XX (AXRL/) AX R L.
XX
XX Zhang HM, Ax RL;
XX
XX WPI; 2002-066307/09.
XX
XX P-PSDB; AAU11518.
XX
XX Novel human or bovine fertility-associated antigen useful for
XX stabilising sperm cell acrosome and increasing fertility of a male
XX
XX Claim 1; Fig 2A-E; 54pp; English.
XX
XX The invention describes a novel isolated human fertility-associated
XX antigen (FAA) and a bovine FFA polypeptide. Human or bovine FFA is useful
XX for increasing the stability of a sperm cell acrosome and for increasing
XX the fertility of a human male by administering FFA. Detecting single
XX nucleotide polymorphisms (SNPs) of the human or bovine FFA gene is useful
XX for assaying fertility by detecting the presence or absence of a FFA SNP
XX in the mammal and correlating the frequency of the FFA SNP to the
XX specification. The FFA SNPs are chosen from FFA SNPs 1-34 given in the
XX CC
XX CC fitness of a human or bovine, by determining the nucleotide sequence of
XX CC the FFA SNP, quantifying the fertility of more than one mammal containing
XX CC the FFA SNP and correlating the frequency of the FFA SNP to the
XX CC reproductive fitness of the mammals. FFA improves the integrity of sperm
XX CC membranes and increases the capacitation of sperm derived from either
XX CC fertile or infertile humans. Early genetic identification of infertility
XX CC improves the prognosis for subsequent attempts at fertilisation and
XX CC facilitates early intervention to determine whether the individual may
XX CC benefit from fertility treatments to avoid costly and/or emotional
XX CC problems with attempted inseminations. This sequence encodes the human
XX CC fertility-associated antigen (FAA), HC2, one of 4 partial prostate CDNA
XX CC clones of the novel FFA gene described in the method of the invention.
XX
XX Sequence 350 BP; 88 A; 91 C; 93 G; 78 T; 0 other;
XX
XX Query Match 51.9%; Score 348.4; DB 24; Length 350;
XX Best Local Similarity 99.7%; Pred. No. 5.1e-100;
XX Matches 349; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 203 CAGGAAACCCCTTGTGCTGCTTCCAGTACCTACACCGCTGCAAGAGACTTGAT 262
XX 1 CAGGAAACCCCTTGTGCTGCTTCCAGTACCTACACCGCTGCAAGAGACTTGAT 60
XX
XX 263 TGTCCCTCCGACACACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGAT 322
XX 61 TGTCCCTCCGACACACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGAT 120
XX
XX 323 CTACACAGATGTGAAGCTCGCTGGAATGACAGAAATTCATTTCATGGGACTTCAA 382
XX 121 CTACACAGATGTGAAGCTCGCTGGAATGACAGAAATTCATTTCATGGGACTTCAA 180
XX
XX 383 TGTCTGGCTGCAAGCTGCTCCCAAGAGGCTGGAAGGACATCGCTGAGAGCGGACCC 442
XX 181 TGTCTGGCTGCAAGCTGCTCCCAAGAGGCTGGAAGGACATCGCTGAGAGCGGACCC 240

```

```

OY 443 CAAGTCTGTTGGCTGATCGGGGACCAAGAGACACCGCTCAAGAGACCAACTG 502
DB 241 CAAGTCTGTTGGCTGATCGGGGACCAAGAGACACCGCTCAAGAGACCAACTG 300
OY 503 CGCCTATGACAGATCGTGTCTTAGAGGCAAAATATTGTCAACTCTGATG 552
DB 301 CGCCTATGACAGATCGTGTCTTAGAGGCAAAATATTGTCAACTCTGATG 350

RESULT 14
AA517903
ID AA517903 standard; cDNA; 350 BP.
XX
XX AA517903;
XX
XX 26-MAR-2002 (first entry)
XX
XX partial cDNA encoding human fertility-associated antigen (FAA), HC3.
XX
XX Fertility-associated antigen, FAA; antifertility; sperm cell acrosome;
XX reproductive fitness; single nucleotide polymorphism, SNP; human;
XX HC3; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 2..349
XX /*tag= a
XX /product= "Human FFA"
XX /note= "Human fertility associated antigen, HC3"
XX /partial
XX /note= "No start or stop codon given"
XX /transl_except= (pos:251..253, aa: Arg)
XX
XX MO200176529-A2.
XX
XX 18-OCT-2001.
XX
XX 09-APR-2001; 2001WO-US10802.
XX
XX 07-APR-2000; 2000US-195225P.
XX
XX (ZHAN/) ZHANG H M.
XX (AXRL/) AX R L.
XX
XX Zhang HM, Ax RL;
XX
XX WPI; 2002-066307/09.
XX P-PSDB; AAU11518.
XX
XX Novel human or bovine fertility-associated antigen useful for
XX stabilising sperm cell acrosome and increasing fertility of a male
XX
XX Claim 1; Fig 2A-E; 54pp; English.
XX
XX The invention describes a novel isolated human fertility-associated
XX antigen (FAA) and a bovine FFA polypeptide. Human or bovine FFA is useful
XX for increasing the stability of a sperm cell acrosome and for increasing
XX the fertility of a human male by administering FFA. Detecting single
XX nucleotide polymorphisms (SNPs) of the human or bovine FFA gene is useful
XX for assaying fertility by detecting the presence or absence of a FFA SNP
XX in the mammal and correlating the presence or absence with the fertility
XX of the mammal. The FFA SNPs are chosen from FFA SNPs 1-34 given in the
XX specification. The FFA SNPs are useful for determining the reproductive
XX fitness of a human or bovine, by determining the nucleotide sequence of
XX the FFA SNP, quantifying the fertility of more than one mammal containing
XX the FFA SNP and correlating the frequency of the FFA SNP to the
XX reproductive fitness of the mammals. FFA improves the integrity of sperm
XX membranes and increases the capacitation of sperm derived from either
XX fertile or infertile humans. Early genetic identification of infertility
XX improves the prognosis for subsequent attempts at fertilisation and
XX facilitates early intervention to determine whether the individual may
XX benefit from fertility treatments to avoid costly and/or emotional

```

CC problems with attempted inseminations. This sequence encodes the human  
CC fertility-associated antigen (FAA), HC3, one of 4 partial prostate cDNA  
CC clones of the novel FAA gene described in the method of the invention.  
XX  
SQ Sequence 350 BP; 89 A; 93 C; 94 G; 74 T; 0 other;

Query Match 46.9%; Score 314.8; DB 24; Length 350;  
Best Local Similarity 93.7%; Pred. No. 2.3e-89;  
Matches 328; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 203 CAGGGAACCTTTGTGCTGCTGTTCCAGTCACCTACACCGCTGTCAAGGACTTCGTGAT 262  
DB 1 CAGGAGACCTTTGTGCTGCTGTTCCAAATCTCCACACTGCTGTCAAGACTTCGTGAT 60  
QY 263 TGTCCCTCCGACACACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGCTGATGT 322  
DB 61 TATCCCTCCGACACACCCAGAGACATCCGTTAAGGAGATCGATGAGTTGTTGAGGT 120  
QY 323 CTACACAGATGTGAACGCTGCTGGAATCGAGAGATTTTCAATTCATGGGTGACTTCAA 382  
DB 121 CTACACGAGCTGAACACCGCTGGAAGCGGAGAAATTTCAATTCATGGGTGACTTCAA 180  
QY 383 TGTGCTGCTGAGTACGCTCCCAAGAGCGCTGGAAGGACATCCGCTGAGGACGACCC 442  
DB 181 TGTGCTGCTGAGTACGCTCCCAAGAGCGCTGGAAGGACATCCGCTGAGGACGACCC 240  
QY 443 CAAGTTCTGTTGCTGATCGGGACCAAGAGGACACACCGGTCAAGAGAGACACAACTG 502  
DB 241 CAAGTTCTGTTGCTGATCGGGACCAAGAGGACACACCGGTCAAGAGAGACACAACTG 300  
QY 503 CGCCTATGACAGATCGTCTGTAGAGGACAAATATTTGCAACTCTGGTG 552  
DB 301 CGCCTATGACAGATCGTCTGTAGAGGACAAATATTTGCAACTCTGGTG 350

RESULT 15  
AAS17904  
ID AAS17904 standard; cDNA; 349 BP.  
AC AAS17904;  
XX  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Partial cDNA encoding human fertility-associated antigen (FAA), HC4.  
XX  
XX Fertility-associated antigen; FAA; antiinfertility; sperm cell acrosome;  
KW reproductive fitness; single nucleotide polymorphism; SNP; human;  
KW HC4; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 2..348  
FT /\*tag= a  
FT /product= "Human FAA"  
FT /note= "Human fertility-associated antigen, HC4"  
FT /partial  
FT /note= "No start or stop codon given"  
FT /transl\_except= (pos: 11..12, aa:Phe)  
FT /note= "This codon has an apparent 1 nucleotide  
FT deletion which alters the reading frame"  
XX  
PN W0200176529-A2.  
XX  
PD 18-OCT-2001.  
XX  
PF 09-APR-2001; 2001WO-US10802.  
XX  
PR 07-APR-2000; 2000US-195225P.  
XX  
PA (ZHAN/) ZHANG H M.  
PA (AXRL/) AX R L.  
XX

PI Zhang HM, Ax RL;  
XX WPI; 2002-066307/09.  
DR P-PSDB; AAU11520.  
XX  
XX  
PT Novel human or bovine fertility-associated antigen useful for  
PT stabilising sperm cell acrosome and increasing fertility of a male  
XX  
PS Claim 1; Fig 2A-E; 54pp; English.  
XX  
CC The invention describes a novel isolated human fertility-associated  
CC antigen (FAA) and a bovine FAA polypeptide. Human or bovine FAA is useful  
CC for increasing the stability of a sperm cell acrosome and for increasing  
CC the fertility of a human male by administering FAA. Detecting single  
CC nucleotide polymorphisms (SNPs) of the human or bovine FAA gene is useful  
CC for assaying fertility by detecting the presence or absence of a FAA SNP  
CC in the mammal and correlating the presence or absence with the fertility  
CC of the mammal. The FAA SNPs are chosen from FAA SNPs 1-34 given in the  
CC specification. FAA SNP is also useful for determining the reproductive  
CC fitness of a human or bovine, by determining the nucleotide sequence of  
CC the FAA SNP, quantifying the fertility of more than one mammal containing  
CC the FAA SNP and correlating the frequency of the FAA SNP to the  
CC reproductive fitness of the mammals. FAA improves the integrity of sperm  
CC membranes and increases the capacitation of sperm derived from either  
CC fertile or infertile humans. Early genetic identification of infertility  
CC improves the prognosis for subsequent attempts at fertilisation and  
CC facilitates early intervention to determine whether the individual may  
CC benefit from fertility treatments to avoid costly and/or emotional  
CC problems with attempted inseminations. This sequence encodes the human  
CC fertility-associated antigen (FAA), HC4, one of 4 partial prostate cDNA  
XX clones of the novel FAA gene described in the method of the invention.  
SQ Sequence 349 BP; 89 A; 91 C; 93 G; 76 T; 0 other;

Query Match 44.8%; Score 300.5; DB 24; Length 349;  
Best Local Similarity 91.6%; Pred. No. 7.3e-85;  
Matches 318; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 206 GGAACCCCTTTGTGCTGCTGTTCCAGTCACCTACACCGCTGTCAAGGACTTCGTGATTC 265  
DB 3 GGGAGCCCTTTGTGCTGCTGTTCCATCTCTCCACACTGCTGTCAAGACTTCGTGATTA 62  
QY 266 CCCCTGTCACACCCCTGAGACATCCGTTAGAGAGATTGATGAGCTGGCTGATGTCTA 325  
DB 63 CCCCTGTCACACCCAGAGACATCCGTTAAGGAGATCGATGAGTTGGTTGAGGTCTA 122  
QY 326 CACAGATGTGAACGCTGCTGGAATGCGAGAGAAATTTTCAATTCATGGGTGACTTCAATGC 385  
DB 123 CACGGACGTGAACACCCGCTGGAAGCGGAGAAATTTTCAATTCATGGGTGACTTCAATGC 182  
QY 386 TGGCTGTCAGCTACGTCCCAAGAGCGCTGGAAGGACATCCGCTGAGGACGACCCCAA 445  
DB 183 CGGCTGTCAGCTACGTCCCAAGAGCGCTGGAAGAGAAATTCGCTTGAGGACTGACCCGAG 242  
QY 446 GTTCGTTTGGCTGATCGGGACCAAGAGGACACCAACCGTCAAGAGAGACACAACTTCGCG 505  
DB 243 GTTCTGTTGGCTGATCGGGACCAAGAGGACACCAACCGTCAAGAGAGACACAACTTCGCG 302  
QY 506 CTATGACAGGATCGTGTCTAGAGGACAAAATTTGTCAACTCTCTGGTG 552  
DB 303 CTATGACAGGATCGTGTCTAGAGGACAAAATTTGTCAACTCTCTGGTG 349

Search completed: December 9, 2002, 09:18:03  
Job time : 201.384 secs

**This Page Blank (uspto)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 09:20:03 ; Search time 1496.6 Seconds  
(without alignments)  
7261.235 Million cell updates/sec

Title: US-09-905-114-3

Perfect score: 671

Sequence: 1 acaacagatgtgcccata.....atcatcatgaagaaccatga 671

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_nam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	528.6	78.8	890	9	AL546857
2	526.6	78.5	896	9	AL546432
C 3	518.8	77.3	931	9	AL571894
C 4	511	76.2	833	9	AL572037
C 5	463	69.0	506	10	AV607154
6	458	68.3	666	10	BE048177

C 7	432.6	64.5	642	10	AW543356
C 8	428.8	63.9	695	13	B1522352
C 9	425.4	63.4	825	12	BG677772
10	422.4	63.0	825	13	BG530136
11	415.4	61.9	628	13	B1460950
C 12	412.2	61.4	620	10	AW543802
C 13	407.8	60.8	784	9	A1927844
C 14	407	60.7	610	10	AW540389
15	406.4	60.6	668	10	AW915564
16	402	59.9	832	13	B1838188
C 17	387.6	57.8	634	9	A1660448
C 18	387	57.7	685	13	B1833952
19	377	56.2	759	12	BG534290
C 20	374.2	55.8	566	10	AW538998
21	363.4	54.2	981	14	BQ710632
C 22	352.4	52.5	613	10	AW779563
C 23	351.4	52.4	578	10	AW658627
C 24	351.2	52.3	527	10	AW539589
C 25	335.2	50.0	643	10	AW001760
C 26	329.8	49.2	714	12	BG566893
C 27	321	47.8	542	12	BF434855
28	320.8	47.8	855	12	BF237166
C 29	314.2	46.8	554	10	AW243180
C 30	313	46.6	555	9	A1434430
31	309.8	46.2	901	12	BF303508
C 32	309.2	46.1	487	9	A1194886
C 33	303.8	45.3	497	9	AA564572
34	291.2	43.4	586	13	BM489658
35	284.2	42.4	595	12	BG077563
C 36	283.8	42.3	433	12	BF911384
37	279.8	41.7	341	10	AW803005
C 38	276.6	41.2	511	9	A1271665
C 39	272.6	40.6	523	9	A1248904
40	267.6	39.9	522	14	BQ186387
C 41	264.6	39.4	536	9	AA707043
C 42	263.8	39.3	486	9	AA989198
43	262.6	39.1	453	14	H78170
44	260.4	38.8	427	10	AW827162
C 45	251.8	37.5	404	13	BM386137

#### ALIGNMENTS

RESULT 1  
AL546857  
LOCUS AL546857 890 bp mRNA linear EST 16-FEB-2001  
DEFINITION AL546857 LTI\_NFL006\_PL2 Homo sapiens cDNA clone CSODI026YK08 5  
prime, mRNA sequence.  
ACCESSION AL546857  
VERSION AL546857.1 GI:12880381  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 890)  
AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES  
source location/Qualifiers  
1..890  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CSODI026YK08"  
/clone\_lib="LTI\_NFL006\_PL2"  
/tissue\_type="placenta"  
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT)-primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL : <http://fulllength.invitrogen.com>

BASE COUNT 250 a 215 c 220 g 204 t 1 others  
ORIGIN

Query Match 78.8%; Score 528.6; DB 9; Length 890;  
Best Local Similarity 87.9%; Pred. No. 4.3e-141;  
Matches 576; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

```
QY 1 ACAACAGATCTGCCCATCTAGTATGAGAAAGCTTAAACGAAATTCAGAAAGGCATTA 60
DB 190 ACAACAGATCTGCCCATCTAGTATGAGAAAGCTTAAACGAAATTCAGAAAGGCATTA 249
QY 61 CATACACTATGTATGTAGCTCTCGCTTTGGAAGAAACATATTAAGAACATATGCTT 120
DB 250 CATACACTATGTATGTAGCTCTCGCTTTGGAAGAAACATATTAAGAACATATGCTT 309
QY 121 TTCTCTATTAAGAAAGCTAGTGTCTGTAAACAAAGCTACTACACGACTATCAG 180
DB 310 TTCTCTATTAAGAAAGCTAGTGTCTGTAAACAAAGCTACTACACGACTATCAG 369
QY 181 CTGAGACGCGAGATGTGTTTTCAGAGAAACCTTTGTGCTGTGTTCCAGTACCTTCA 240
DB 370 ATGAGACGCGAGATGTGTTTTCAGAGAAACCTTTGTGCTGTGTTCCAGTACCTTCA 429
QY 241 CCGCTGTCAAGACTTCTGATGTTCCCTTGACACACACCCCTGAGCATCCGTTAAG 300
DB 430 CTGCTGTCAAGACTTCTGATGTTCCCTTGACACACACCCCTGAGCATCCGTTAAG 489
QY 301 AGATTGATGAGCTGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGAGAAATT 360
DB 480 AGATTGATGAGCTGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGAGAAATT 549
QY 361 TCATTTTCATGAGGTGACTTCAATGCTGTGCTGAGCTACGCTCCCAAGAGGCTTGAAG 420
DB 550 TCATTTTCATGAGGTGACTTCAATGCTGTGCTGAGCTACGCTCCCAAGAGGCTTGAAG 609
QY 421 ACATCCGCTGAGAGCGAAGCCCAAGTTCGTTGGCTGATCGGGAGCAAGAGGACACCA 480
DB 610 ACATCCGCTGAGAGCGAAGCCCAAGTTCGTTGGCTGATCGGGAGCAAGAGGACACCA 669
QY 481 CGGTCAAGAAAGACCAAACTGCGCTATGACAGATCGTGTAGAGGACAAATATTG 540
DB 670 CGGTCAAGAAAGACCAAACTGCGCTATGACAGATCGTGTAGAGGACAAATATTG 729
QY 541 TCAACTGTGTGTCTCTCAATCAAACTCGTCTTGTGATTTCCAGAAAGCTTACAGTTGT 600
DB 730 TCAACTGTGTGTCTCTCAATCAAACTCGTCTTGTGATTTCCAGAAAGCTTACAGTTGT 789
QY 601 CTGAATCGAAGGCCCTGATGTCAGGACCACTTCCAGTTTCATCATCATCATCA 655
DB 790 CTGAATCGAAGGCCCTGATGTCAGGACCACTTCCAGTTTCATCATCATCATCA 844
```

RESULT 2  
AL546432 896 bp mRNA linear EST 16-FEB-2001  
LOCUS AL546432 LTI\_NFL006.PL2 Homo sapiens cDNA clone CS0D1030YD20 5  
DEFINITION prime, mRNA sequence.  
ACCESSION AL546432  
VERSION AL546432.1 GI:12879540  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 896)

AUTHORS Li, W.B., Gruber, C., Joesse, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr).

FEATURES  
source  
1. 896  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0D1030YD20"  
/clone\_1b="LTI\_NFL006.PL2"  
/tissue\_type="placenta"  
/note="Vector: pCMVSPORT 6, Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL : <http://fulllength.invitrogen.com>"

BASE COUNT 249 a 217 c 226 g 203 t 1 others  
ORIGIN

Query Match 78.5%; Score 526.6; DB 9; Length 896;  
Best Local Similarity 87.6%; Pred. No. 1.6e-140;  
Matches 574; Conservative 1; Mismatches 80; Indels 0; Gaps 0;

```
QY 1 ACAACAGATCTGCCCATCTAGTATGAGAAAGCTTAAACGAAATTCAGAAAGGCATTA 60
DB 240 ACAACAGATCTGCCCATCTAGTATGAGAAAGCTTAAACGAAATTCAGAAAGGCATTA 299
QY 61 CATACACTATGTATGTAGCTCTCGCTTTGGAAGAAACATATTAAGAACATATGCTT 120
DB 300 CATACACTATGTATGTAGCTCTCGCTTTGGAAGAAACATATTAAGAACATATGCTT 359
QY 121 TTCTCTATTAAGAAAGCTAGTGTCTGTAAACAAAGCTACTACACGACTATCAG 180
DB 360 TTCTCTATTAAGAAAGCTAGTGTCTGTAAACAAAGCTACTACACGACTATCAG 419
QY 181 CTGAGACGCGAGATGTGTTTTCAGGAAACCTTTGTGCTGTGTTCCAGTACCTTCA 240
DB 420 CTGAGACGCGAGATGTGTTTTCAGGAAACCTTTGTGCTGTGTTCCAGTACCTTCA 479
QY 241 CCGCTGTCAAGACTTCTGATGTTCCCTTGACACACACCCCTGAGCATCCGTTAAG 300
DB 480 CTGCTGTCAAGACTTCTGATGTTCCCTTGACACACACCCCTGAGCATCCGTTAAG 539
QY 301 AGATTGATGAGCTGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGAGAAATT 360
DB 540 AGATTGATGAGCTGCTGATGTCTACACAGATGTGAAACGTCGCTGGAATGAGAAATT 599
QY 361 TCATTTTCATGAGGTGACTTCAATGCTGTGCTGAGCTACGCTCCCAAGAGGCTTGAAG 420
DB 600 TCATTTTCATGAGGTGACTTCAATGCTGTGCTGAGCTACGCTCCCAAGAGGCTTGAAG 659
QY 421 ACATCCGCTGAGAGCGAAGCCCAAGTTCGTTGGCTGATCGGGAGCAAGAGGACACCA 480
DB 660 ACATCCGCTGAGAGCGAAGCCCAAGTTCGTTGGCTGATCGGGAGCAAGAGGACACCA 719
QY 481 CGGTCAAGAAAGACCAAACTGCGCTATGACAGATCGTGTAGAGGACAAATATTG 540
DB 720 CGGTCAAGAAAGACCAAACTGCGCTATGACAGATCGTGTAGAGGACAAATATTG 779
QY 541 TCAACTGTGTGTCTCTCAATCAAACTCGTCTTGTGATTTCCAGAAAGCTTACAGTTGT 600
DB 780 TCAACTGTGTGTCTCTCAATCAAACTCGTCTTGTGATTTCCAGAAAGCTTACAGTTGT 839
QY 601 CTGAATCGAAGGCCCTGATGTCAGGACCACTTCCAGTTTCATCATCATCATCA 655
```





RESULT	5
LOCUS	AV607154/c
DEFINITION	AV607154 506 bp mRNA linear EST 28-NOV-2001 AV607154 Bos taurus kidney fetus Bos taurus cDNA clone EIKI039B05
ACCESSION	3'', RNA sequence.
VERSION	AV607154
KEYWORDS	AV607154.1 GI:9737527 EST.
SOURCE	COW.
ORGANISM	Bos taurus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
AUTHORS	Bovidae; Bovinae; Bos. 1 (bases 1 to 506) Takasuga,A., Hirotsune,S., Itoh,R., Jitchazono,A., Suzuki,H., Aso,H. and Sugimoto,Y.
TITLE	Establishment of a high throughput EST sequencing system using poly(A) tail-removed cDNA libraries and determination of 36,000 bovine ESTs
JOURNAL	Nucleic Acids Res. 29 (22), EI08 (2001)
MEDLINE	21570584
COMMENT	Contact: Yoshikazu Sugimoto Animal Genetics Division Shirakawa Institute of Animal Genetics OdaKura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan Tel.: 81-248-25-5641 Fax: 81-248-25-5725 Email: kazusugi@cocoa.ocn.ne.jp This clone was obtained from a polyA-deleted cDNA library.

[illegible]

```
Seq primer: -40RP from Gibco
High quality sequence stop: 431.
Location/Qualifiers
1. .666
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2291681"
/clone_lib="NCI_CGAP_Brn52"
/tissue_type="tumor, 5 pooled (see description)"
/lab_host="DH10B"
/Note="Organ: brain; Vector: pCMV-SPORT6; Site 1: Sali;
Site 2: NotI; This library represents the normalized
version of NCI CGAP_Brn35. Cloned unidirectionally.
Primer: Oligo dr. Average insert size 1.19 kb. Tumor
types include: meningioma, oligodendroglioma, astrocytoma
(grade II), medulloblastoma, astrocytoma (grade IV).
Constructed by Life Technologies."
BASE COUNT 198 a 149 c 170 g 147 t 2 others
ORIGIN

Query Match 68.3%; Score 458; DB 10; Length 666;
Best Local Similarity 88.1%; Pred. No. 8e-121;
Matches 497; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 1 ACAACAGGATCTGCCCATCTACTGATGGAAGCTAAACGGAATTCACAGAAAGGCATAA 60
Db 95 ACAACAGGATCTGCCCATCTACTGATGGAAGCTGAACAGAAATTCACAGGAGGCATAA 154

QY 61 CATACAACTATGTGATTAGCTCTCGCCTTGGAGAAACACATATAAAGAACATATGCCT 120
Db 155 CGTACAACTATGTGATTAGCTCTCGCCTTGGAGAAACACATATAAAGAACATATGCCT 214

QY 121 TTCTCTATAAGAAAGCTAGTGTCTGTAAAAACAAAGCTTACCTTACCAACGACTATCAGG 180
Db 215 TTCTCTATAAGAAAGCTAGTGTCTGTAAAAACAAAGCTTACCTTACCAACGACTATCAGG 274

QY 181 CTGGAGACCCAGATGTTTCCAGGAAACCCCTTGTGTCTGTGTTCCAGTCAACCTACA 240
Db 275 ATGGAGACCCAGATGTTTCCAGGAAACCCCTTGTGTCTGTGTTCCAGTCAACCTACA 334

QY 241 CGCTGTCAAGGACTTCGTGATTGTCCTCCCTGCACACACCCCTCGAGACATCGTTAGG 300
Db 335 CTGCTGTCAAGGACTTCGTGATTGTCCTCCCTGCACACACCCCTCGAGACATCGTTAGG 394

QY 301 AGATTGATGAGCTGGCTGATGTCTACACAGATGTGAACCGTCTGCTGGAATGCAGAAATT 360
Db 395 AGATCGATGAGTGGTGTGAGTCTACACGAGCTGAACACCGCTGGAAGCGGAGATT 454

QY 361 TCATTTTCATGGTGACTTCATGCTGGCTGAGCTACGCTCCCAAGAGGCTCGAAGG 420
Db 455 TCATTTTCATGGTGACTTCATGCTGGCTGAGCTACGCTCCCAAGAGGCTCGAAGG 514

QY 421 ACATCCGCTGAGGACGACCCAGTTCGTTGGCTGATCGGGACCAAGAGGACACCA 480
Db 515 ACATCCGCTGAGGACGACCCAGTTCGTTGGCTGATCGGGACCAAGAGGACACCA 574

QY 481 CGGTCAAGAGAGCAACAACTCGCTATGACGAGGATCGTCTTAGAGGACAAATATTG 540
Db 575 CGGTCAAGAGAGCAACAACTCGTATGATGACGAGGATCGTCTTAGAGGACAAATATC 634

QY 541 TCAACTCTGGTGGTCTCTCAATCAA 564
Db 635 GCAGTTCGTGTGTCCCAAGTCAA 658

RESULT 7
AW543356/c 642 bp mRNA linear EST 31-AUG-2000
LOCUS C0168H07-3 NIA Mouse E7.5 Extraembryonic Portion cDNA Library Mus
DEFINITION musculus cDNA clone C0168H07 3', mRNA sequence.
ACCESSION AW543356
VERSION AW543356.1 GI:7185773
KEYWORDS EST.
```

```
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 642)
AUTHORS Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grahovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H., Wood,W.H., III, Becker,K.G. and KO,M.S.H.
TITLE Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
MEDLINE 203811348
COMMENT Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: C0168 row: H column: 07
Seq primer: -21M13 Forward
High quality sequence stop: 642
POLYA=Yes.
Location/Qualifiers
1. 642
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="niaEST:C0168H07-3"
/db_xref="taxon:10090"
/clone="C0168H07"
/clone_lib="NIA Mouse E7.5 Extraembryonic Portion cDNA Library"
/dev_stage="7.5dpc Embryo"
/lab_host="DH10B"
/Note="vector: pSPORT1 (Gibco/BRL Life Technology);
Site 1: Sali; Site 2: NotI; Total RNAs were extracted from
5 EPC. The double-stranded cDNA was synthesized by
Gibco's kit with an Oligo(dT) primer (NotI primer-adaptor
from GibcoBRL)
[5'-pGACTAGTCTAGATCGGAGCGGCCCTTTTTCCTTTT-3']
from 0.8ug of mRNA. The double-stranded cDNAs were
treated with T4 DNA polymerase and purified by
ethanol-precipitation. The cDNAs were ligated to
Lone-linker LL-Sal3 (include Sali sequence). The cDNAs
were purified by phenol/chloroform and separated from
free linkers by Centricon 100. Then, cDNAs were amplified
by long-range high fidelity PCR using Takara's Ex Taq
polymerase. Then, the cDNAs were purified by
phenol/chloroform and by Centricon 100. The cDNAs were
digested with Sali and NotI enzymes. Then, the cDNAs were
size selected by Gibco's Size Fractionation Column. The
cDNAs were cloned into Sali/NotI site of pSPORT1 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by chemical method. The library was
constructed by Xiaohong Wang."
BASE COUNT 166 a 161 c 148 g 167 t
ORIGIN

Query Match 64.5%; Score 432.6; DB 10; Length 642;
Best Local Similarity 82.6%; Pred. No. 1.6e-113;
Matches 495; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 57 ATACATACAACTATGTGATTAGCTCTCGCCTTGGAGAAACACATATAAAGACAGTAT 116
Db 642 ACACATACAACTATGTGATTAGCTCTCGCCTTGGAGAAACACATATAAAGACAGTAT 583

QY 117 GCCTTTCTATAAGAAAGCTAGTGTCTGTAAACAAAGCTACCTTACCAACAGTAT 176
Db 582 GCCTTCGTCTACAAGGAGAGCTGTGTCTGTGAAGACAAAATACCACTACCATGACTAT 523

QY 177 CAGGCTGGAGAGCAGATGTCTTTTCCAGGAAACCCCTTTGGTCTGGTTCAGTCAACC 236
Db 522 CAGGATGGAGACACAGACGCTTTTCCAGGAGCCCTTTGTGGTTTGGTTCATTCCTCCC 463
```

QY 237 TACACCGCTGTCAAGACCTTGTGATTTGCCCTGACACCAACCCCTGAGACATCCGTT 296  
 DB 462 TTTACTGCTGTCAAGACCTTGTGATTTGCCCTGACACCAACCCCTGAGACATCCGTT 403  
 QY 297 AGAGAGATGATGAGCTGGGTGATGCTCAACAGATGTGAAAAGTGTGAGTGAAGAG 356  
 DB 402 AAGAGATGATGAGCTGGGTGATGCTCAACAGATGTGAAAAGTGTGAGTGAAGAG 343  
 QY 357 AATTGATTTTCAATGGGTAATCTTCAATGCTGCTGACAGTACGTCCTCCCAAGAGCCTG 416  
 DB 342 AATTGATTTTCAATGGGTAATCTTCAATGCTGCTGACAGTACGTCCTCCCAAGAGCCTG 283  
 QY 417 AAGGATATCCGCTGAG 476  
 DB 282 CAGAAATTTGCTTGAAG 223  
 QY 477 ACCACGCTGAG 536  
 DB 222 ACTACGCTGAG 163  
 QY 537 ATTGTCATCTGTGGTGTCTTCAATCAATCAATCAATCAATCAATCAATCAATCAAT 596  
 DB 162 ATAGTCAATCTGTGGTGTCTTCAATCAATCAATCAATCAATCAATCAATCAATCAAT 103  
 QY 597 TTGCTGATTCGAAGGCTCTGATGTCACGACCACTTTCCAGTTTATCATCATCA 655  
 DB 102 TTGCTGATTCGAAGGCTCTGATGTCACGACCACTTTCCAGTTTATCATCATCA 44

RESULT 8  
 B1522352/c 695 bp mRNA linear EST 29-AUG-2001  
 LOCUS 60308142011 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:522052 3'  
 DEFINITION mRNA sequence.  
 ACCESSION B1522352  
 VERSION B1522352.1 GI:15347144  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 695)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LLM11554 row: h column: 05  
 High quality sequence start: 18  
 High quality sequence stop: 695.  
 Location/Qualifiers  
 1..695

FEATURES  
 Source  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5220533"  
 /clone\_lib="NIH\_MGC\_120"  
 /lab\_host="DH10B"  
 /note="Organ: pooled pancreas and spleen; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male; Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH\_MGC Library."

BASE COUNT 159 a 171 c 163 g 202 t  
 ORIGIN  
 Query Match 63.9%; Score 428.8; DB 13; Length 695;  
 Best Local Similarity 87.2%; Pred. No. 2.1e-112;  
 Matches 516; Conservative 0; Mismatches 72; Indels 4; Gaps 4;

QY 49 GAAAAGGATTAACATATCAATATGATTAAGTCTGCTTGGAGAAACATATAAG 108  
 DB 695 GGAGAGCATTAACATATGATTAAGTCTGCTTGGAGAAACATATAAG 636  
 QY 109 AACATATGCTTTCTCTATAAAGAAAGCTAGTCTGTAAACAACTACTTACC 168  
 DB 635 CACAAATATGCTTTCTCTATAAAGAAAGCTAGTCTGTAAACAACTACTTACC 576  
 QY 169 ACAGCTATCAGGCTGAG 228  
 DB 575 ATGACTATACAGATGAGAACCCAGATGTGTTTCCAGGAGACCTTTGTGATCTGATCC 516  
 QY 229 AGTCAACCTTAC -CCGCTGTCAAGAGACTTGTGATTTGCCCTGACACCAACCCCTGAG 287  
 DB 515 AATCTCCCAATCTGTCTGTCAGAGACTTGTGATTTGCCCTGACACCAACCCCTGAG 456  
 QY 288 ACATCCCTTAAAGAGATGATGAGCTGCT -GATGCTACACAGATGTGAACCTGCTG 346  
 DB 455 ACATCCCTTAAAGAGATGATGAGCTGCT -GATGCTACACAGATGTGAACCTGCTG 396  
 QY 347 GAATCAAGAGATTTTCAATTTTCAATGAGTCAATGCTGCTGACCTGACCTGCTGCTG 406  
 DB 395 GAAGGCGAGAAATTTTCAATTTTCAATGAGTCAATGCTGCTGACCTGACCTGCTGCTG 336  
 QY 407 GAAGGCTTGAAGAGACATCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 466  
 DB 335 GAAGGCTTGAAGAGACATCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 276  
 QY 467 CCAAGAGAGACCAACGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 526  
 DB 275 CCAGAGAGACCAACGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 216  
 QY 527 AGGACAAATATTTTCAATCTGCTGAGTCTCAATCA -AACCTGCTTTGATTTCCAGA 585  
 DB 215 AGGACAAATATTTTCAATCTGCTGAGTCTCAATCA -AACCTGCTTTGATTTCCAGA 156  
 QY 586 AAGCTTAAGTGTGCTGAATCAAGGCTGAGTGTGAGGACCACTTCC 637  
 DB 155 AAGCTTAAGTGTGCTGAATCAAGGCTGAGTGTGAGGACCACTTATTC 105

RESULT 9  
 LOCUS BG687772 825 bp mRNA linear EST 29-MAY-2001  
 DEFINITION 602786883F1 NCI\_CGAP\_SG2 Mus musculus cDNA clone IMAGE:4912893 5',  
 mRNA sequence.  
 ACCESSION BG687772  
 VERSION BG687772.1 GI:14218312  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 1 (bases 1 to 825)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov

Plate: LLAM10817 row: e column: 22  
High quality sequence stop: 799.  
Location/Qualifiers  
1. . 825  
/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4912893"  
/clone\_lib="NCI CGAP SG2"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:  
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo  
dT. Average insert size 1.3 kb. Constructed by Life  
Technologies. Note: this is a NCI CGAP Library."  
BASE COUNT 215 a 203 c 204 g 203 t  
ORIGIN  
Query Match 63.4%; Score 425.4; DB 12; Length 825;  
Best Local Similarity 82.7%; Pred. No. 2.2e-111; Indels 3; Gaps 3;  
Matches 522; Conservative 0; Mismatches 106; Indels 3; Gaps 3;  
QY 2 CAACAGGATCTGCCCATACTGATGGAGAGCTTAAACGGAAATTCAGAAAGGCATAAC 61  
DB 196 CAACACATCTGCCCATCTGATGGAGAGCTGATGGAAATTCAGAAAGGCACAAAT 255  
QY 62 ATACAACTATGATTTAGTCTTCGCTTGGAGAAACACATATAAAGAACAGTATGCTTT 121  
DB 256 ATACAACTATGATTTAGTCTTCGACTTGGAGAAACACGTTACAAAGAGCAGTATGCTTT 315  
QY 122 TCTCTATAAGAAAGCTAGTCTGTCTTAACAAAGCTACCTTACACGACTATCAGGC 181  
DB 316 CGTGTACAAGGAGAGCTGGTGTCTGTGAAGACAAATATACCTACCTGACTATCAGGA 375  
QY 182 TGGAGACGAGATGTGTTTTCAGGAAACCTTTGTGGTCTGTGCTCCAGTCAACCTTACAC 241  
DB 376 TGGAGACACAGCGTGTTCAGGAGAGCCCTTTGTGGTTCATTCCTCCCTTTAC 435  
QY 242 CGTGTCAAGGACTTCGTGATTTGCCCTTGACACACCCCTTGAGACATCCGTTTAGAGA 301  
DB 436 TGTGTCAAGGACTTCGTGATTTGCCCTTGCACAACTCCCGAGACCTCCGTTAAAGA 495  
QY 302 GATTGATGAGCTGGCTGATGTCTACACAGATGGAACGTGCTGGAATGCAGAGAAATT 361  
DB 496 GATGATGAGCTGGCTGATGTCTACACGATGTGAAGCCAGTGGAGACAGAGAAATT 555  
QY 362 CATTTTCATGGGTGACTTCAATGCTGGCTGCAGTACGTCCCAAGAGCCCTGGGAAGA 421  
DB 556 CATCTTCATGGGTGATTTCAACGCGGCTGTAGCTATGTCCCAAGAGCCCTGGCAGAA 615  
QY 422 CATCCGCTGA -GGACGGACCCCAAGTTCCG -TTTGGCTGATCGGGACCAAGAGCACAC 479  
DB 616 CATTCGTTTGAAGGACGGACCCCAAGTTTGTGTTGGCTGATTTGGGACCAAGAGGACAC 675  
QY 480 ACGGTCAAGAGACACAACTGCGGCTATGACAGATCGTGTAGAGACAAATATT 539  
DB 676 ACGGTCAAGAGAGTACACAGCTGTGCTATGACAGATTTGTGTTTGTGGCAAGAGATA 735  
QY 540 GTCAACTCTGGTGGTCTCAATCAAACTCGTCTTTGATTTCCAGAAAGCTTACAGTTG 599  
DB 736 GTCAACTCTGGTGGTCTCCCGTTCAGTGGCGTCTTTGACTTTTCAGAAAGCTTATGACT 795  
QY 600 TCTGAATCGAAGGCCCTGGATGTACAGCAC 630  
DB 796 TCTGAGCG -AGAGCCCTGGATGTACGGATC 825  
RESULT 10  
BG530136  
LOCUS 825 bp mRNA linear EST 03-APR-2001  
DEFINITION 602558626F1 NIH\_MGC\_61 Homo sapiens cDNA clone IMAGE:4696638 5',  
mRNA sequence.  
ACCESSION BG530136  
VERSION BG530136.1 GI:13521673

KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS 1 (bases 1 to 825)  
TITLE NIH-MGC http://mgi.nci.nih.gov/.  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: CLONETECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM1524 row: c column: 07  
High quality sequence stop: 743.  
Location/Qualifiers  
1. . 825  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4696638"  
/clone\_lib="NIH\_MGC\_61"  
/tissue\_type="embryonal carcinoma"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site\_1:  
SfiI (ggccattatggcc); Site\_2: SfiI (ggccattatggcc);  
Double-stranded cDNA was prepared from cell line RNA. 5'  
and 3' adaptors were used in cloning as follows: 5'  
adaptor sequence: 5'-CACGGCATTATGGCC-3' and 3' adaptor  
sequence: 5'-ATTCTAGAGCCGAGCGGCGACATG-dt(30)BN-3'  
(where B = A, C, or G and N = A, C, G, or T). Average  
insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies  
contained inserts by PCR. This library was enriched for  
full-length clones and was constructed by Clontech  
Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC  
Library."  
BASE COUNT 225 a 204 c 217 g 179 t  
ORIGIN  
Query Match 63.0%; Score 422.4; DB 12; Length 825;  
Best Local Similarity 83.2%; Pred. No. 1.6e-110;  
Matches 516; Conservative 0; Mismatches 101; Indels 3; Gaps 3;  
QY 1 ACAACAGGATCTGCCCATCTGATGGAGAGCTTAAACGGAAATTCAGAAAGGCATAA 60  
DB 206 ACAACAGGATCTGCCCATCTGATGGAGAGCTTAAACGGAAATTCAGAGAGGCATAA 265  
QY 61 CATACAACTATGTGATTTAGCTCTCGCTTGGAGAAACACATATAAAGAACAGTATGCCT 120  
DB 266 CCGTCAACTATGTGATTTAGCTCTCGCTTGGAGAAACACATATAAAGAACAAATATGCCT 325  
QY 121 TTCTCTATAAGAAAGCTAGTGTCTGTAAAAACAAAGCTACCTTACCACGACTATCAGG 180  
DB 326 TTCTCTACAGAAAGCTGGTGTCTGTCAAGAGAGGTATCACTACCATGACTATCAGG 385  
QY 181 CTGG -AGACGCGAGATGTGTTTCCAGGAAACCCCTTTGTGGTCTGTTCCAGTACCCCTAC 239  
DB 386 ATGGCAGACGCGAGATGTGTTTCCAGGAGCCCTTTGTGGTCTGTTCCAAATCTCCCCAC 445  
QY 240 ACCGCTGTCAAGGACTTCGTGATTTGCTCCCTGACACACCCCTTGAGACATCCGTTAGA 299  
DB 446 ACTGCTGTCAAGAGCTTCGTGATTTATCCCCCTTGACACACAGCCCGAGACATCCGTTAG 505  
QY 300 GAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAACACGCTCGCTGGAATGCAGAGAT 359  
DB 506 GAGATCGATGATGTTGAGGTCTACACGAGAGCTGAAACACCGCTTGGAGGCGGGAAT 565  
QY 360 TTCATTTTCATGGGTGACTTCAATGCTCGCTGACGTACGTCCTCCCAAGAGGCGCTGGAA - 418

DB	Accession	Source	Organism	Reference	Authors	Title	Journal	Comment
Db	566	TTGATTTTATGAGGTCGACTTCAATGCGCGCTGAGCTTGACGTCCCAAGAAGGCTGTGAAC	625					
QY	419	GGACATCCGCGCTGAGAGACGGACCCCAAGTTCGTTTGGCGATGCGGGGACCAAGAGACAC	478					
Db	626	GAAACATCCGCTTGAGGACGTGACCCCGAGTTTGTTGGCTGATCGGGGACCAAGAGAGACAC	685					
QY	479	CACGGTCAGAAAGAGACCAAACTGCCCTCATGACAGGA-TCCGTGCTTAGAGAGCAAAATA	537					
Db	686	CACGGTCAGAAAGAGACCACTGTCATATATGACAGAGTTGGTGCTTAGAGAGCAAAATA	745					
QY	538	TTGTCACTCTGTGGTCTCTCAATCAAACCTGCTTTGATTTTCCAGAAAGCTTACAGGT	597					
Db	746	AATCGTCACGTCTGTGCTCCCAAGTCAAAACAGGTCCTTGACTTCCAGAAAGCTTCCAGGC	805					
QY	598	TGTCGATCAAGAGGCGCTG 617						
Db	806	TGACTGTAGAGAGGCGCTTG 825						
RESULT 11								
LOCUS	BI460950	628 bp	mRNA	linear	EST 21-AUG-2001			
DEFINITION	603207187P1 NIH_MGC_97 Homo sapiens CDNA clone IMAGE:5272917 5',							
ACCESSION	BI460950							
VERSION	BI460950.1	GI:15251606						
KEYWORDS	EST.							
SOURCE	human.							
ORGANISM	Homo sapiens							
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
AUTHORS	Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.							
TITLE	1 (bases 1 to 628)							
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.							
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)							
	Unpublished (1999)							
	Contact: Robert Strusberg, Ph.D.							
	Email: cgapbs-remail.nih.gov							
	Tissue Procurement: Miklos Palkovits, M.D., Ph.D.							
	CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki							
	Toshiyuki and Piero Carninci (RIKEN)							
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)							
	DNA Sequencing by: Incyte Genomics, Inc.							
	Clone distribution: MGC clone distribution information can be							
	found through the I.M.A.G.E. Consortium/LLNL at:							
	http://image.llnl.gov							
	Plate: LLM11688 row: n column: 22							
	High quality sequence stop: 628.							
FEATURES	Location/Qualifiers							
source	1..628							
	/organism="Homo sapiens"							
	/db_xref="taxon:9606"							
	/clone="IMAGE:5272917"							
	/clone_11b="NIH_MGC_97"							
	/lab_host="DH10B"							
	/note="Organ: testis; Vector: pBluescriptR (modified							
	pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gcgag							
	); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',							
	size-selected for ROT 5. This is a primary library enriched							
	for full-length clones and constructed using the							
	Cap-trapper method (Carninci, in preparation). Library							
	constructed by M. Brownstein (NHGRI/NHGR), National							
	Institutes of Health). Note: this is a NIH_MGC Library."							
BASE COUNT	181 a 144 c 166 g 137 t							
ORIGIN								
Query Match	61.9%	Score 415.4	DB 13	Length 628				
Best Local Similarity	90.6%	Pred. No. 1.5e-108						
Matches 443	Conservative 0	Mismatches 46	Indels 0	Gaps 0				
QY	1	ACAACAGCATCTGCCCATCTGATGAGAGAACTAAACGAAATTCAGAAAGGCATTA 60						
Db	140							

OY	61	CATACAACTATGTATGTAGCTCTCGSCCTTGGAAAGAAACACATATPAAAGAACGATGTGCT	120
Db	200	CATACAACTATGTATGTATGTAGCTCTCGSCCTTGGAAAGAAACACATATPAAAGAACATATGCTT	259
OY	121	TTCTCTATPAAAGAAAGCTAGTGTCTGTAAAACAAAGCTACCTTACCAAGCATATCAGG	180
Db	260	TTCTCTAACAGGAAAGCTAGTGTCTGTAAAGAGAGATTATCATCATCATGACTATCAGG	319
OY	181	CTGAGACGACAGATGTGTCTTCCAGAGGAAACCTTGTGTGTCTGTGTCCAGTCAACCTTCA	240
Db	320	ATGGAGACGCGAGATGTGTCTTCCAGAGAGCCCTTGTGTGTGTGTCCAAATCTCCCA	379
OY	241	CCGCTGTCAAGGACTTCTGTGATTGTCCCCCTTGCAACACACCCTTGAGACATCCGTTAGAG	300
Db	380	CTGCTGTCAAGACTTCTGTGATTATCCCCCTTGCAACACACCCTTGAGACATCCGTTAAGG	439
OY	301	AGATTGATGAGCTGGCTGTATGTCTTACACAGATGTGAAACGTGTGTGAAATGCAGAAATT	360
Db	440	AGATTGATGAGTGTGTGTAGGTCTTACAGGACGTGAAACACCGCTGGAAGCGGGAATT	499
OY	361	TCATTTTATGAGGTACTTCAATGTGTGCTCAGCTACGTCCCCAAGAAAGGCTGTGAAGG	420
Db	500	TCATTTTATGAGGTACTTCAATGTGTGCTCAGCTACGTCCCCAAGAAAGGCTGTGAAGG	559
OY	421	ACATCCGCTTAAAGACGAGACCCCAAGTTCGTTGGCTGATCGGGGACCAAGAGACACCA	480
Db	560	ACATCCGCTTAAAGACTACCCCAAGTTCGTTGGCTGATCGGGGACCAAGAGGACACCA	619
OY	481	CGGTCAAGA 489	
Db	620	CGGTCAAGA 628	

RESULT 12	
AM543802/c	
LOCUS	620 bp mRNA linear EST 31-AUG-2000
DEFINITION	musculus CDNA clone C0174C07 3', mRNA sequence.
ACCESSION	AM543802
VERSION	AM543802.1 GI:7186219
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
AUTHORS	1 (bases 1 to 620) Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grabovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagataja,R., Doi,H., Wood,W.H., Ili, Becker,K.G. and Ko,M.S.H.
TITLE	Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental CDNA microarray
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
MEDLINE	20381148
COMMENT	Contact: George J. Kargul Laboratory of Genetics National Institute on Aging/National Institutes of Health 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA Email: cdna@gsun.grc.nia.nih.gov Plate: C0174 row: C column: 07 Seq primer: -21M13 Forward High quality sequence stop: 620 POLYA=Yes.
FEATURES	
SOURCE	Location/Qualifiers 1..620 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="niaEST:C0174C07-3" /db_xref="taxon:10090" /clone="C0174C07" /clone_1lb="NIA Mouse E7.5 Extraembryonic Portion cDNA Library" /sex="unknown"

RESULT 13  
AI927844/C

BASE COUNT	173 a	188 c	179 g	242 t	2 others
ORIGIN	Fatima Bonaïdo. "				
Query Match	60.8%; Score 407.8; DB 9; Length 784;				
Best Local Similarity	86.1%; Pred. No. 2.5e-106;				
Matches 451; Conservative	0; Mismatches 73; Indels 0; Gaps 0;				
Qy	132	GAAGAAGCTAGTGTCTGTGAACAAAGCTACCTCTACCACGACTATCAGGCTGGAGACGCA	191		
Db	667	GAAGAAGCTAGTGTCTGTGAAGAGGAGTTATCTACCATGACTATCAGGATGGAGACGCA	608		
Qy	192	GATGTGTTTTCCAGGGAAACCCCTTGTGGTCTGGTTCAGTTCACCCCTACACCGCTGTCAAG	251		
Db	607	GATGTGTTTTCCAGGGAGCCCTGTGTGTCTGGTTCCAATCTCCACACACTGCTGTCAAA	548		
Qy	252	GACTTCGTGATTTGTCCCTCTGCACACACCCCTGAGACATCCGTTTAGAGAGATTTGATGAG	311		
Db	547	GACTTCGTGATTTATCCCTCTGCACACCCACNCAGAGACATCCCGTTAAGGAGATCGATGAG	488		
Qy	312	CTGGCTGATGCTCTACACAGATGTGAACGTCGCTGGNAATGCAGAGAAATTTTCATTTTCATG	371		
Db	487	TTGGTTGAGGCTCTACCGGACGTGAACACCCGCTGGAAGCGGAGAAATTTTCATTTTCATG	428		
Qy	372	GGTGACTTTCAAATGTGSGCTSCAGCTAGTCCCAAGAGGCCCTGGGAAGGACATCCCGCTG	431		
Db	427	GGTGACTTTCAAATCGCGCTCGACGCTAGTCCCCAAGAGGCCCTGGGAAGAACATCGCTTG	368		









**This Page Blank (uspto)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 08:59:53 ; Search time 40.9082 Seconds  
(without alignments)  
5030.295 Million cell updates/sec

Title: US-09-905-114-3

Perfect score: 671

Sequence: 1 acaacaggatctgcccata.....atcatcatgaagaaccatga 671

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/prodata/1/ina/5A-COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5B-COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A-COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B-COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/PCTUS-COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	481.2	71.7	1208	1	US-08-640-765A-2
2	481.2	71.7	1208	3	US-09-073-613-2
3	163.2	24.3	1039	1	US-07-895-300A-2
4	163.2	24.3	1039	5	PCT-US93-05136-2
5	163.2	24.3	1039	5	PCT-US93-05136-2
6	163.2	24.3	1039	5	PCT-US93-10519-1
7	161.6	24.1	1039	1	US-08-458-367-2
8	124.6	18.6	1615	2	US-08-491-988-15
9	124.6	18.6	1624	2	US-08-491-988-14
10	124.6	18.6	1648	2	US-08-491-988-13
11	119	17.7	1055	2	US-08-468-012A-1
12	119	17.7	1055	4	US-09-054-989-1
13	35.2	5.2	1427	3	US-09-173-581-10
14	35.2	5.2	1427	4	US-08-420-915-10
15	32.8	4.9	1293	2	US-08-924-440-1
16	32.2	4.8	289	4	US-09-007-005-17
17	32.2	4.8	289	4	US-09-244-796-17
18	32.2	4.8	7218	1	US-08-232-463-14
19	31.8	4.7	1511	2	US-08-809-763-3
20	31.8	4.7	1511	3	US-08-956-253-3
21	31.6	4.7	1001	4	US-09-641-638-90
22	31.2	4.6	248	3	US-08-513-974B-300
23	31.2	4.6	248	3	US-08-513-974B-301
24	30.6	4.6	1920	1	US-08-487-753-1
25	30.6	4.6	1920	2	US-08-480-065-1
26	30.6	4.6	1920	3	US-08-487-744-1
27	30.6	4.6	1920	5	PCT-US93-09167-1

C 28	30.6	4.6	2900	2	US-09-027-337-9	Sequence 9, Appli
C 29	30.6	4.6	2900	4	US-09-644-600-9	Sequence 9, Appli
C 30	30.6	4.6	3147	2	US-09-027-337-1	Sequence 1, Appli
C 31	30.6	4.6	3147	4	US-09-644-600-1	Sequence 1, Appli
C 32	30.6	4.6	3147	4	US-09-644-600-18	Sequence 18, Appli
C 33	30	4.5	1734	4	US-09-484-970B-63	Sequence 63, Appli
C 34	30	4.5	2071	1	US-07-923-724-1	Sequence 1, Appli
C 35	30	4.5	2071	2	US-08-609-426A-1	Sequence 1, Appli
C 36	30	4.5	2071	2	US-08-374-652C-3	Sequence 3, Appli
C 37	30	4.5	4049	1	US-08-162-809-17	Sequence 17, Appli
C 38	30	4.5	4097	1	US-08-162-809-11	Sequence 11, Appli
C 39	29.8	4.4	3981	2	US-08-955-138-2	Sequence 2, Appli
C 40	29.6	4.4	47	4	US-09-146-054-7	Sequence 7, Appli
C 41	29.6	4.4	47	4	US-09-664-977A-7	Sequence 7, Appli
C 42	29.6	4.4	270	4	US-09-146-054-8	Sequence 8, Appli
C 43	29.6	4.4	270	4	US-09-664-977A-8	Sequence 8, Appli
C 44	29.6	4.4	352	4	US-09-228-986-21	Sequence 21, Appli
C 45	29.6	4.4	560	3	US-09-042-426-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1  
US-08-640-765A-2  
; Sequence 2, Application US/08640765A  
; Patent No. 5821103  
; GENERAL INFORMATION:  
; APPLICANT: Tanuma, Sei-ichi  
; TITLE OF INVENTION: NOVEL DEOXYRIBONUCLEASE  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Leydig, Voit & Mayer  
; STREET: Two Prudential Plaza, Suite 4900  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60601  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/640,765A  
; FILING DATE: 06-MAY-1996  
; CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6239518  
; FILING DATE: 06-SEP-1994  
; CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
; NAME: Green, Robert F.  
; REGISTRATION NUMBER: 27555  
; REFERENCE/DOCKET NUMBER: 73362  
TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 616-5600  
; TELEFAX: (312) 616-5700  
INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1208 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 12..941  
US-08-640-765A-2

Query Match 71.7%; Score 481.2; DB 1;  
Best Local Similarity 85.1%; Pred. NO. 5.8e-139;  
Matches 550; Conservative 0; Mismatches 93; Indels 3; Gaps 1;

OY 1 ACAACAGATCTGCCCTACTGATGAGAGAGCTAAACGGAATTCAGAAAAAGCATTA 60  
DB 217 ACAACAAATCTGTCTCCAGCTGATGAGAGAGCTGATGAGAAATCAACAAAGACAGA 276  
OY 61 CATACAACTATGATTTAGCTCTCGCTTGAAGAAACATATTAAGAACATATGCTT 120  
DB 277 CATACAACTATGATTTAGCTCTCGCTTGAAGAAACATATTAAGAACATATGCTT 336  
OY 121 TTCTCTATTAAGAAAAAGTAGTGTCTGTAACAAAGCTACCTCTACACGATATCAG 180  
DB 337 TCCTCTACAAAGAGAGAGCTGTGTCTGTGAAGCAAAAATACCTACCTACCTATCAG 396  
OY 181 CTGAGAGAGCAATATGTTTTCAGAGAACCTTTGTGTGTGTGTGTGTGTGTGTGTGT 240  
DB 397 ATGAGAGACACAGAGGTTTTCAGAGAACCTTTGTGTGTGTGTGTGTGTGTGTGTGT 456  
OY 241 CCGCTGTCAAGAGCTTGT 300  
DB 457 CTGCTGCCAAGAGACTTGT 516  
OY 301 AGATTGATGAGCTGT 360  
DB 517 AGATGATGAGCTGT 576  
OY 361 TCATTTCATGAGT 420  
DB 577 TCATTTCATGAGT 636  
OY 421 ACATCCGCTGTGAG 480  
DB 637 ACATCCGCTGTGAG 696  
OY 481 CGGTCAAG 540  
DB 697 CGGTCAAG 756  
OY 541 TCAACTCTGT 600  
DB 757 TCAACTCTGT 816  
OY 601 CTGAATCGAAGGCGCTG--GATGTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 643  
DB 817 CTGAAGAGAGAGGCGCTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 862

RESULT 2  
US-09-073-613-2  
; Sequence 2, Application US/09073613  
; Patent No. 6143875  
; GENERAL INFORMATION:  
; APPLICANT: Tanuma, Sei-ichi  
; TITLE OF INVENTION: ANTIBODY TO NOVEL DEOXYRIBONUCLEASE  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Leydig, Voit & Mayer  
; STREET: Two Prudential Plaza, Suite 4900  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/073.613  
; FILING DATE: 06-MAY-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/640,765  
; FILING DATE:

APPLICATION NUMBER: JP 6239518  
FILING DATE: 06-SEP-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: LARCHER, Carol  
REGISTRATION NUMBER: 35243  
REFERENCE/DOCKET NUMBER: 85519  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 616-5600  
TELEFAX: (312) 616-5700  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1208 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 12..941  
US-09-073-613-2

Query Match 71.7%; Score 481.2; DB 3; Length 1208;  
Best Local Similarity 85.1%; Pred. No. 5.8e-139;  
Matches 550; Conservative 0; Mismatches 93; Indels 3; Gaps 1;

OY 1 ACAACAGATCTGCCCTACTGATGAGAGAGCTAAACGGAATTCAGAAAAAGCATTA 60  
DB 217 ACAACAAATCTGTCTCCAGCTGATGAGAGAGCTGATGAGAAATCAACAAAGACAGA 276  
OY 61 CATACAACTATGATTTAGCTCTCGCTTGAAGAAACATATTAAGAACATATGCTT 120  
DB 277 CATACAACTATGATTTAGCTCTCGCTTGAAGAAACATATTAAGAACATATGCTT 336  
OY 121 TTCTCTATTAAGAAAAAGTAGTGTCTGTAACAAAGCTACCTCTACACGATATCAG 180  
DB 337 TCCTCTACAAAGAGAGAGCTGTGTCTGTGAAGCAAAAATACCTACCTACCTATCAG 396  
OY 181 CTGAGAGAGCAATATGTTTTCAGAGAACCTTTGTGTGTGTGTGTGTGTGTGTGTGT 240  
DB 397 ATGAGAGACACAGAGGTTTTCAGAGAACCTTTGTGTGTGTGTGTGTGTGTGTGTGT 456  
OY 241 CCGCTGTCAAGAGCTTGT 300  
DB 457 CTGCTGCCAAGAGACTTGT 516  
OY 301 AGATTGATGAGCTGT 360  
DB 517 AGATGATGAGCTGT 576  
OY 361 TCATTTCATGAGT 420  
DB 577 TCATTTCATGAGT 636  
OY 421 ACATCCGCTGTGAG 480  
DB 637 ACATCCGCTGTGAG 696  
OY 481 CGGTCAAG 540  
DB 697 CGGTCAAG 756  
OY 541 TCAACTCTGT 600  
DB 757 TCAACTCTGT 816  
OY 601 CTGAATCGAAGGCGCTG--GATGTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 643  
DB 817 CTGAAGAGAGGCGCTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 862

RESULT 3  
US-07-895-300A-2  
; Sequence 2, Application US/07895300A

```

; Patent NO. 5279823
;
; GENERAL INFORMATION:
;
; APPLICANT: Frenz, John
;
; APPLICANT: Shire, Steven J.
;
; APPLICANT: Sliwkowski, Mary B.
;
; TITLE OF INVENTION: PURIFIED FORMS OF DNase
;
; NUMBER OF SEQUENCES: 17
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: Genentech, Inc.
;
; STREET: 460 Point San Bruno Blvd
;
; CITY: South San Francisco
;
; STATE: California
;
; COUNTRY: USA
;
; ZIP: 94080
;

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/895_300A
FILING DATE: 19920608
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Hensley, Max D.
REGISTRATION NUMBER: 27,043
REFERENCE/DOCKET NUMBER: 747
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1039 bases
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
US-07-895-300A-2

```

Query Match	24.3%	Score 163.2	DB 1	Length 1039
Best Local Similarity	57.4%	Pred. No. 9.5e-41		
Matches 335	Conservative 0	Mismatches 243	Indels 6	Gaps
60	ACATACAACTATGTGATTAGCTCTCGCCTTTGGAAAGAACACATATAAAGAACAGTAGTATGCC	119		
QY				
DB				
409	ACCTATCACTACGTGGTCAGTGAGCCACTGGGACGAAACAGCTATTAAGAGCGCTACCTG	468		
QY				
DB				
120	TTTCTCTATAAGAAAAAGCTAGTGTCTGTAAAAACAAAGCTACTCTTACCACGA---CTAT	176		
QY				
DB				
469	TTCGTGTACAGGCTCACCAGGTGTCTGCGGTGCACAGCTACTACTACGATGATGCGTCG	528		
QY				
DB				
177	CAGGCTGGAGACCCAGATGTGTTTTCCAGGGAAACCCCTTTGTGCTGTGTTCCAGTCAACC	236		
QY				
DB				
529	GAGCCCTCGGGAAACACACCTTCAACCGAGAGCCAGCCATTGTCAGGTTCTTCTCCCGG	588		
QY				
DB				
237	TACACCGCTGTCAAGGACTTCGTGATTGTCCCCCTGCACACCAACCCCTCAGAGCATCCGTT	296		
QY				
DB				
589	TTCACAGAGGTCAAGGAGTTTGCCATTGTTCCCTCGATCGCGCCCGGGGACGCAGTA	648		
QY				
DB				
297	AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATTGTGAACGTCGCTGGAAATGCAGAG	356		
QY				
DB				
649	GCCGAGATCGACGCTCTCTATGACGTCCTACCTGGATGTCCAAGAGAAATGGGGCTTTGGAG	708		
QY				
DB				
357	AATTTTCATTTTCATGGGTGACTTCAATGTGCGCTGCAGCTACGTCCTCCCAAGAGGCGCTGG	416		
QY				
DB				
709	GAGCTCATGTTGATGGCGNCTTCAATGCGGCTGCAGCTATGTGAGACCTCTCCAGTGG	768		
QY				
DB				
417	AAGGACATCCGCTGAGGACGGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC	476		
QY				
DB				
769	TCATCATCCGCTGTGGACAAGCCACCTTCCAGTGGCTGATCCCCACAGCGCTGAC	828		
QY				
DB				

Qy	477	ACACGGTCAAGAGAGACA	CAAACTGCGCCTATGACAGGATCGTCTTAGAGGACAAAAT	536
Db	829	ACCACAG---CTACACCCACGCACTGTGCCTATGACAGGATCGTGGTTGCAGGATCGCTG	885	
Qy	537	ATTGTCAACTCTGGTGGTCTCTCAATCAAACCTCGTCTTTGATTTTCAGAAAGCTTACAGG	596	
Db	886	CTCCGAGGCGCGTGTGTTCCGCACTCGGCTCTTCCCTTTAACTTCAGGCTGCCTATGGC	945	
Qy	597	TTGTCTGAATCGAAGGCCCTGGATGTGACGACCACTTTCCAGT	640	
Db	946	CTGAGTCACCAACTGGGCCCAAGCCATCAGTGACCACTATCCAGT	989	

```

RESULT 4
US-09-638-112-2
; Sequence 2, Application US/09638112
; Patent No. 6440412
; GENERAL INFORMATION:
; APPLICANT: Frenz, John
; APPLICANT: Shire, Steven J.
; APPLICANT: Slikowski, Mary B.
; TITLE OF INVENTION: PURIFIED FORMS OF DNase
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/638,112
; FILING DATE: 09-Aug-2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/942561
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Johnston, Sean A.
; REGISTRATION NUMBER: 35,910
; REFERENCE/DOCKET NUMBER: P0747C7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3562
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1039 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-09-638-112-2

```

Query Match	24.3%;	Score 163.2;	DB 4;	Length 1039;
Best Local Similarity	57.4%;	Pred. No. 9.5e-41;		
Matches 335;	Conservative 0;	Mismatches 243;	Indels 6;	Gaps
Qy	60	ACATACAACTATGTGATTAGCTCTCGCCTTGGAGAAACACATATATAAGAACAGTAGTGCC	119	
Db	409	ACCTATCACTACGTGGTGCAGTGAGCCACTGGGACGGACAGCTATTAAAGAGCGCCTACCTG	468	
Qy	120	TTTCTCTATAAAGAAAAGCTAGTGTCTGTGTAAAAACAAAGCTTACCTCTACCACGAA---CTAT	176	
Db	469	TTCGTGTACAGGCCTGACCAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGCGTGC	528	
Qy	177	CAGCGTGGAGACGAGATGTGTTTTCCAGGAAACCCCTTTGTGTGTCTGGTTCAGTCACCC	236	
Db	529	GAGCCCTTCGGGAACGACACCTTCAACGAGAGCCAGCCATTCTCAGTTCTTCTCTCCGG	588	

QY 237 TACACCGCTGTCAAGAGACTTCTGTATGTCCTCCCTGACACACCCCTTGAGACATCCGTT 296  
 Db 589 TTCAAGAGAGTCAAGAGAGATTGTCATGTTCCCTGACAGCGCCCGGGGAGACGACAGTA 648  
 QY 237 AGAGAGATTGATGAGCTGGCTGATGCTCAACAGATGTGAAAGCTGCGTAATGACAGAG 356  
 Db 649 GCCGAGATCGAGCTCTCTATGACGCTTACCTGAGATGTCAGAAAGAAATGAGGCTTGGAG 708  
 QY 357 AATTCAATTTTCATGAGGAGTCTTCAATGCTGAGTCACTAGCTCCCAAGAGGCTTGG 416  
 Db 709 GACGTCATGTTGATGAGGAGCACTTCAATGAGGAGCTGAGCTATGAGACCTCCAGTGG 768  
 QY 417 AAGACATCCGCTGAG 476  
 Db 769 TCATCCATCCGCTGTGAG 828  
 QY 477 ACCACGTCAG 536  
 Db 829 ACCACAG--CTACACCCAG 885  
 QY 537 ATTGTCAACTCTGTGTCCTTCAATCAAACTGCTTTGATTTTCAGAAAGCTTACAGG 596  
 Db 886 CTCGAGAGCGCGTGTGTCCTGAGCTGAGCTTCCCTTTAACTTCAAGGCTGCTATGAGC 945  
 QY 597 TTGTCTGATGAG 640  
 Db 946 CTGAGTGACCAACTGCGCCCAAGCAGATCAGTGACCACTATCCAGT 989

## RESULT 5

PCT-US93-05136-2

Sequence 2, Application PC/TUS9305136

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.

TITLE OF INVENTION: PURIFIED FORMS OF DNase

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: patin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/05136

FILING DATE: 19930528

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Johnston, Sean A.

REGISTRATION NUMBER: P35,910

REFERENCE/DOCKET NUMBER: 747PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-3562

TELEFAX: 415/952-9881

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1039 bases

TYPE: NUCLEIC ACID

STRANDEDNESS: single

TOPOLOGY: linear

PCT-US93-05136-2

Query Match

24.3%, Score 163.2, DB 5, Length 1039;

Best Local Similarity 57.4%; Pred. No. 9.5e-41;  
 Matches 335; Conservative 0; Mismatches 243; Indels 6; Gaps 2;  
 QY 60 ACATACCACTGTGATGATTAGCTCTGCTTGGAGAAACATATTAAGACAGTATGCC 119  
 Db 409 AACTATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 468  
 QY 120 TTTCTCTATAAAGAAAGCTATGCTCTGTATAAACAAGCTATCTTACACAGA--CTAT 176  
 Db 469 TTGCTGATCAGGCGTCAAGAGTGTCTGCGTGACAGACTACTACTAGATGAGGCTGC 528  
 QY 177 CAGGCTGAGAGAGAGATGATGTTTCCAGGGAACCTTTGGCTGCTGCTGCTGCTGCTG 236  
 Db 529 GAGCCCTCGGGAACGACACTTCAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 588  
 QY 237 TACACCGCTGTCAAGAGACTTGTGATGTTCCCTGACACACCCCTGAGACATCCGTT 296  
 Db 589 TTCAAGAGAGTCAAGAGATTGGCATTTGCTTCCCTGATGCGGCCCGGGGAGAGAGAGTA 648  
 QY 297 AGAGAGATTGATGAGCTGCTGATGCTCAACAGATGTGAACGTCCTGGAATGACAGAG 356  
 Db 649 GCCGAGATCGAGCTCTCTATGACGCTACCTGATGTCCAAGAGAAATGGGAGCTTGGAG 708  
 QY 357 AATTCAATTTTCATGAGGAGATCAATGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 416  
 Db 709 GACGTCATGTTGATGAGGAGATTTCAATGCGGCTGCTGAGCTATGAGAGAGAGAGAGAG 768  
 QY 417 AAGACATCCGCTGAG 476  
 Db 769 TCATCCATCCGCTGTGAG 828  
 QY 477 ACCACGTCAG 536  
 Db 829 ACCACAG--CTACACCCAG 885  
 QY 537 ATTGTCAACTCTGTGTCCTTCAATCAAACTGCTTTGATTTTCAGAAAGCTTACAGG 596  
 Db 886 CTCGAGAGCGCGTGTGTCCTGAGCTGAGCTTCCCTTTAACTTCAAGGCTGCTATGAGC 945  
 QY 597 TTGTCTGATGAG 640  
 Db 946 CTGAGTGACCAACTGCGCCCAAGCAGATCAGTGACCACTATCCAGT 989

## RESULT 6

PCT-US93-10519-1

Sequence 1, Application PC/TUS9310519

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.

TITLE OF INVENTION: Compaction Assay for Assessment of Respiratory Disease

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: patin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/10519

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Johnston, Sean A.

```
; REGISTRATION NUMBER: 35,910
; REFERENCE/DOCKET NUMBER: 792C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3562
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1039 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
PCT-US93-10519-1

Query Match      24.3%; Score 163.2; DB 5; Length 1039;
Best Local Similarity 57.4%; Pred. No. 9.5e-41;
Matches 335; Conservative 0; Mismatches 243; Indels 6; Gaps 2;

QY 60 ACATACACTATGTGATTAGCTCTCGCTTGGGAAGAAACACATATAAAGAACAGTATGCC 119
Db 409 ACCTATCACTACGTGTGCTAGTACGACCTGGGACGGAACAGCTATAAGGAGCGCTACCTG 468
QY 120 TTCTCTCTATAAAGAAAGCTAGTGTCTGTAACAAAGCTACCTTACCACGA---CTAT 176
Db 469 TTCGTGTACAGGCTCACCAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGCTGC 528
QY 177 CAGGCTGGAGACGAGATGTGTTTCCAGGGAACCCCTTGTGGTCTGGTTCAGTACCCC 236
Db 529 GAGCCCTCGGGAACGACACCTTCAACCGAGAGCCAGCCATTGTCTCTCCCG 588
QY 237 TACACGCTGTCAAGGACTTTCGTGATGTCCTGACAGCTCGTGTCTAGAGCAAAAT 536
Db 829 ACACAG---CTACACCCAGCGACTGTGCTATGACAGGATCGTGTGAGGGATGCTG 885
QY 537 ATTGTCAACTCTGGTGGTCTCAATCAAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 596
Db 886 CTCGAGGCGCGTGTGTTCCCGACTCGGCTCTTCCCTTTAACTTCCAGGCTGCCTATGCG 945
QY 597 TTGCTCTGAATCGAAGCGCCCTGGATGTCAGCGACCACTTTTCCAGT 640
Db 946 CTGAGTGACCAACTGCGCCCAAGCCATCAGTGACCACTATCCAGT 989
```

## RESULT 7

```
US-08-458-367-2
; Sequence 2, Application US/08458367
; Patent No. 5783433
; GENERAL INFORMATION:
; APPLICANT: Frenz, John
; APPLICANT: Shire, Steven J.
; APPLICANT: Slivkowski, Mary B.
; TITLE OF INVENTION: PURIFIED FORMS OF DNase
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
```

```
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,367
; FILING DATE: 02-Jun-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/409631
; FILING DATE: 22-Mar-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/348284
; FILING DATE: 30-No. 5783433-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/116186
; FILING DATE: 02-Sep-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/895300
; FILING DATE: 08-Jun-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Johnston, Sean A.
; REGISTRATION NUMBER: 35,910
; REFERENCE/DOCKET NUMBER: P0747C4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3562
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1039 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-458-367-2
```

```
Query Match      24.1%; Score 161.6; DB 1; Length 1039;
Best Local Similarity 57.2%; Pred. No. 3e-40;
Matches 334; Conservative 0; Mismatches 244; Indels 6; Gaps 2;

QY 60 ACATACAACTATGTGATTAGCTCTCGCTTGGGAAGAAACACATATAAAGAACAGTATGCC 119
Db 409 ACCTATCACTACGTGTGCTAGTACGACCTGGGACGGAACAGCTATAAGGAGCGCTACCTG 468
QY 120 TTCTCTCTATAAAGAAAGCTAGTGTCTGTAACAAAGCTACCTTACCACGA---CTAT 176
Db 469 TTCGTGTACAGGCTCACCAGGTGTCTGCGGTGGACAGCTACTACTACGATGATGCTGC 528
QY 177 CAGGCTGGAGACGAGATGTGTTTCCAGGGAACCCCTTGTGGTCTGGTTCAGTACCCC 236
Db 529 GAGCCCTCGGGAACGACACCTTCAACCGAGAGCCAGCCATTGTCTCTCCCG 588
QY 237 TACACGCTGTCAAGGACTTTCGTGATGTCCTGACAGCTCGTGTCTAGAGCAAAAT 536
Db 589 TTCAGAGGTCAGGAGTTTCCCAATGTTCCCTCTGATGCGGCCCCGGGGACCGAGTA 648
QY 297 AGAGAGATTGATGAGCTGGCTGATGCTTACACAGATGTGAAACGTCGCTGGAATCGAGAG 356
Db 649 GCCGAGATCGACGCTCTCTATGACGCTTACCTGGATGTCCAAGAGAAATGGGGCTTGGAG 708
QY 357 AATTTCAATTTTCATGGGTGACTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 416
Db 709 GAGTCATGTTGATGGGCACTTCAATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 768
QY 417 AAGGACATCCGCTGAGGACGCAACCCCAAGTTTCGTTGCTGCTGCTGCTGCTGCTGCTG 476
Db 769 TCATCATCCGCTGTGGACAAAGCCCCACCTTCCAGTGGCTGATCCCCGACGCGCTGAC 828
QY 477 ACCACGGTCAAGAGAGGACAAACTGCGCCTATGACAGGATCGTGTGCTTAGAGGACAAAAT 536
```

Db 829 ACCACAG--CTACACCCAGCAGCTGTGCTATGACAGGATGTGGTTCAGGAGTCTG 885  
Qy 537 ATTGTCAACTGTGTGTCTTCATCAAACTGTCTTTGATTTCCAGAAAGCTTACAG 596  
Db 886 CTCGACGCGCGGTGTGTTCCTCCGACTCGCTCTTCCCTTTAATTCCAGGCTCCTATG 945  
Qy 597 TTGTCTGAATCGAAGCCCTGTGATGTACGACCACTTTCCAGT 640  
Db 946 CTGAGTGAACCACTGGCCCAAGCCATCAGTACCACTATCCAGT 989

## RESULT 8

US-08-491-988-15  
Sequence 15, Application US/08491988  
Patent No. 5973116  
GENERAL INFORMATION:  
APPLICANT: EPENETOS, AGAMENON A.  
APPLICANT: SPOONER, ROBERT A.  
APPLICANT: DEONARAIN, MAHENDRA  
TITLE OF INVENTION: Compounds for targeting  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MCAULAY NISSEN GOLDBERG KIEL & HAND, LLP  
STREET: 261 MADISON AVENUE  
CITY: NEW YORK  
STATE: NY  
COUNTRY: USA  
ZIP: 10016-2391  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/491,988  
FILING DATE: 18-DEC-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: GOLDBERG, JULES E.  
REGISTRATION NUMBER: 24,408  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-986-4090  
TELEFAX: 212-818-9479  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1615 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
US-08-491-988-15

Query Match 18.6%; Score 124.6; DB 2; Length 1615;  
Best Local Similarity 53.2%; Pred. No. 1e-28;  
Matches 312; Conservative 0; Mismatches 269; Indels 6; Gaps 2;

Db 60 ACATACAACTATGTGATTAGCTCTGCTTGGAGAAACATATATAAGACAGTATGCC 119  
Qy 1009 ACCATCACTAGTGTGTTTCCGACCCCTCGGGCTTAATCTTACAAAGACGTAAG 1068  
Qy 120 TTTCTCTATAAAGAAAGCTAGTGTCTTAATAAACAAGCTACTTACACAGAC--TAT 176  
Db 1069 TTCCTGTCCGTCGCAAAAGTTTCAGTACTGATACCTACAGTACGACGAGATGC 1128  
Qy 177 CAGGCTGAGACGAGATGTGTTTCCAGGGAACCTTTGTGTGCTGGTCCAGTACACC 236  
Db 1129 GAATTTTCGGTACGACTCTTTCTCCGGGAACCGGCTGTGTTAATTTTCGAGCCAC 1188  
Qy 237 TACACCGCTGTCAAGAGCTTGTGATGTGTCCCTGACACACCCCTGAGACATCCGTT 296  
Db 1189 TCTACCAAGTTAAAGAGTTCGCTATGCTGTGCTGACAGGCGCGCTGAGCGTGT 1248

Qy 297 AGAGAGATTGATGAGCTGCTGATGTCTACACAGATGTGAAGCTCGCTGGAATGACAG 356  
Db 1249 GCTGAATCACTCTCTGTACGAGTGTACCTGAGCGTTGACGAAATGGCACCTGAC 1308  
Qy 357 AATTTCATTTTCAATGGGTGATTTCAATGCTGGCTGACGCTACGCTCCCAAGAGCCTGG 416  
Db 1309 GACGTCATGCAATGGGTGATTTCAAGCTGACGCTCTTATATGATCTTCTTCAGTGG 1368  
Qy 417 AAGGACATCCGCTGAGACGCAACCCAGTTCGTTTGGCTGATCGGAGACCAAGAGAC 476  
Db 1369 TCATGATTCGTCTGCGCACCCTGTCACCTTCCAGTGTGATCCGAGCTCCGCTGAC 1428  
Qy 477 ACCACGGTCAAGAAAGACACAACTGCGCTATGACAGATGCTGTTAGAGACAAAT 536  
Db 1429 ACCACGGTACT--AGTACCACTGCGCTTACGACCGTATGTTGTGCTGATCCCTG 1485  
Qy 537 ATTGTCAACTGTGTGTCTTCATCAAACTGTCTTTGATTTCCAGAAAGCTTACAG 596  
Db 1486 CTGAGTCTCTGTGTGAACCGGTAGCGCGCCCGTTGCACTTCAGGCTGCATATGCT 1545  
Qy 597 TTGTCTGAATCGAAGCCCTGTGATGTACGACCACTTTCCAGTTCA 643  
Db 1546 CTTTCGAACGAATGGCGCTGCGCATCTGTGATCACTACCGGTTGA 1592

## RESULT 9

US-08-491-988-14  
Sequence 14, Application US/08491988  
Patent No. 5973116  
GENERAL INFORMATION:  
APPLICANT: EPENETOS, AGAMENON A.  
APPLICANT: SPOONER, ROBERT A.  
APPLICANT: DEONARAIN, MAHENDRA  
TITLE OF INVENTION: Compounds for targeting  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MCAULAY NISSEN GOLDBERG KIEL & HAND, LLP  
STREET: 261 MADISON AVENUE  
CITY: NEW YORK  
STATE: NY  
COUNTRY: USA  
ZIP: 10016-2391  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/491,988  
FILING DATE: 18-DEC-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: GOLDBERG, JULES E.  
REGISTRATION NUMBER: 24,408  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-986-4090  
TELEFAX: 212-818-9479  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1624 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
US-08-491-988-14

Query Match 18.6%; Score 124.6; DB 2; Length 1624;  
Best Local Similarity 53.2%; Pred. No. 1e-28;  
Matches 312; Conservative 0; Mismatches 269; Indels 6; Gaps 2;

Qy 60 ACATACAACTATGTGATTAGCTCTGCGCTTGGAAAGAAACATATAAAGACAGTATGCC 119

Db 1018 ACCTACACCTACCTGTTTCTGTAACCCCTCGGCGTAACCTTTACAAAGAACGGTACCTG 1077  
Qy 120 TTTCTCTATAAAGAAAGCTAGTGTCTGTAAACAAAGCTACCTTACCAACGAC---TAT 176  
Db 1078 TTTCTGTTCGTCGGAACAAAGTTTCAGTACTGGATACCTACAGTACGACGACGATGC 1137  
Qy 177 CAGGCTGGAGACGAGATGTGTTTTCAGGAAACCTTTGTGTGTCTGTTTCCAGTACACC 236  
Db 1138 GAATCTTGGCGTAACGACTCTTTCTCCGGAACCGGCTGTTGTTAAATTTCTCGAGCCAC 1197  
Qy 237 TACACCGCTGTCAAGACATTCGTAATGTCCTGACACCAACCCCTGAGACATCCGTT 296  
Db 1198 TCTACCAAGTTAAAGATTCGCTATGCTGCTGACACGCGCGCTGCTGACGCTGTT 1257  
Qy 297 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAACGTCGCTGGAATCGAGAG 356  
Db 1258 GCTGAAATCAACTCTCTGTACGAGGTTTACCTGGAGCTTACAGCAGAAATGGCACCTGAC 1317  
Qy 357 AATTTCAATTTCAAGGAGCTTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 416  
Db 1318 GAGCTCATGTGATGCTGACTTCAACGCTGACTGCTCTTATGTAACCTCTTCTCAGTGG 1377  
Qy 417 AAGGACATCGCTGAGGACGACCCCAAGTTCTGTTGCTGTGATCGGGACCAAGAGGAC 476  
Db 1378 TCATCGATTGCTGTGCGCACTCTCGACCTTCCAGTGGCTGATCCCGGACTCCGCTGAC 1437  
Qy 477 ACCACCGTCAAGAAGAGCACAACCTCGCTATGACAGGATCGTGTGTAGAGGACAAAT 536  
Db 1438 ACCACCGCTACT---AGTACCAATCGCTTACGACCGTATGCTGTTGCTGTGATCCCTG 1494  
Qy 537 ATTGTCAACTCTGGTGTCTCAATCAAACTCGTCTTTGATTTCCAGAAAGCTTACAGG 596  
Db 1495 CTGCACTCTTCTGTACCGGTAGCGGCGCCGCTCGACTTCCAGGCTGCATATGTT 1554  
Qy 597 TTGCTGATCGAGCCCTGATGTACGGACCACTTTTCCAGTTCA 643  
Db 1555 CTTTCAACGAAATGCGCTGGCCATCTCTGATCACTACCCCGTTGA 1601

## RESULT 10

US-08-491-988-13  
; Sequence 13, Application US/08491988  
; Patent No. 5973116  
; GENERAL INFORMATION:  
; APPLICANT: EPENETOS, AGAMENNON A.  
; APPLICANT: SPOONER, ROBERT A.  
; APPLICANT: DEONARAIN, MAHENDRA  
; TITLE OF INVENTION: Compounds for targeting  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MCAULAY NISSEN GOLDBERG KIEL & HAND, LLP  
; STREET: 261 MADISON AVENUE  
; CITY: NEW YORK  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10016-2391  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/491,988  
; FILING DATE: 18-DEC-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: GOLDBERG, JULES E.  
; REGISTRATION NUMBER: 24,408  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-986-4090  
; TELEFAX: 212-818-9479  
; INFORMATION FOR SEQ ID NO: 13:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1648 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; US-08-491-988-13

Query Match 18.6%; Score 124.6; DB 2; Length 1648;

Best Local Similarity 53.2%; Pred. No. 1e-28; Indels 6; Gaps 2;  
Matches 312; Conservative 0; Mismatches 269;

Qy 60 ACATACAACATATGTGATTAGCTCTCGCCCTTGAAGAAACACATATAAAGAACAGTATGCC 119  
Db 1042 ACCTACCACTACGTTGTTTCTGAAACCCCTCGGCGTAACCTTTACAAAGAACGGTACCTG 1101  
Qy 120 TTTCTCTATAAAGAAAGCTAGTGTCTGTAAACAAAGCTACCTTACCAACGAC---TAT 176  
Db 1102 TTTCTGTTCGTCGGAACAAAGTTTTCAGTACTGGATACCTACCAAGTACGACGACGATGC 1161  
Qy 177 CAGGCTGGAGACGAGATGTGTTTCCAGGGAACCCCTTGTGTGTGTGTGTGTGTGTGTGTGT 236  
Db 1162 GAATCTTGGCGTAAAGACTCTTCTCCCGGAACCGGCTGTTGTTAAATTTCTGAGCCAC 1221  
Qy 237 TACACCGCTGTCAAGGACTTCGTGATTTGTCCTCCCTGCACACCAACCCCTGAGACATCCGTT 296  
Db 1222 TCTACCAAGTTAAAGAGTTTCGCTATCGTTGCTCTGCACAGCGCGCTCTGACGCTGTT 1281  
Qy 297 AGAGAGATTGATGAGCTGGCTGATGTCTACAGATGTGAACGCTGCTGGAATGACAGAG 356  
Db 1282 GCTGAAATCAACTCTCTGTACGACGTTTACCTGGAGCTTTCAGCAGAAATGGCACCTGAAC 1341  
Qy 357 AATTTCAATTTTCATGGTGACTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 416  
Db 1342 GACGTCATGCTGATGGTGACTTCAACGCTGACTGCTTATGTATTAACCTTCTTCTCAGTGG 1401  
Qy 417 AAGGACATCCGCTGAGGACGACCCCAAGTTTCGTTGCTGATCGGGACCAAGAGGAC 476  
Db 1402 TCATCGATTCTGCTCGGACCTCTGCTGACCTTCCAGTGGCTGATCCCGGACTCCGCTGAC 1461  
Qy 477 ACCAGGCTCAAGAAGAGCACAACCTGCGCTATGACAGGATCGTGTGTAGAGGACAAAAT 536  
Db 1462 ACCACCGCTACT---AGTACCAACTGCGCTTACGACCGTATGTTGCTGCTGATCCCTG 1518  
Qy 537 ATTGTCAACTCTGGTGTCTCTCAATCAAACTCGTCTTTGATTTCCAGAAAGCTTACAGG 596  
Db 1519 CTGCACTCTTCTGTTGATCCGCGGTAGCGCGCCCGCTTCCAGCTTCCAGGCTGCATATGTT 1578  
Qy 597 TTGTCTGAATCAAGGCGCTGATGTACGCGACCACTTTTCCAGTTCA 643  
Db 1579 CTTTCAACGAAATGCGCTGCGCATCTCTGATCACTACCCCGTTGA 1625

## RESULT 11

US-08-468-012A-1  
; Sequence 1, Application US/08468012A  
; Patent No. 5830744  
; GENERAL INFORMATION:  
; APPLICANT: RUBEN, ET AL.  
; TITLE OF INVENTION: Human DNase  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
; ADDRESSEE: CECCHI, STEWART & OLSTEIN  
; STREET: 6 BECKER FARM ROAD  
; CITY: ROSELAND  
; STATE: NEW JERSEY  
; COUNTRY: USA  
; ZIP: 07068  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 INCH DISKETTE  
; COMPUTER: IBM PS/2



```

1  OPERATING SYSTEM:  MS-DOS
2  SOFTWARE:  WORD PERFECT 5.1
3  CURRENT APPLICATION DATA:
4  APPLICATION NUMBER:  US/08/468,012A
5  FILING DATE:  6 JUNE 1995
6  CLASSIFICATION:  435
7  PRIOR APPLICATION DATA:
8  APPLICATION NUMBER:  PCT/US94/04954
9  FILING DATE:  5 MAY 1994
10 ATTORNEY/AGENT INFORMATION:
11 NAME:  MULLINS, J.G.
12 REGISTRATION NUMBER:  33,073
13 REFERENCE/DOCKET NUMBER:  325800-428
14 TELECOMMUNICATION INFORMATION:
15 TELEPHONE:  201-994-1700
16 TELEFAX:  201-994-1744
17 INFORMATION FOR SEQ ID NO:  1:
18     SEQUENCE CHARACTERISTICS:
19     LENGTH:  1055 BASE PAIRS
20     TYPE:  NUCLEIC ACID
21     STRANDEDNESS:  SINGLE
22     TOPOLOGY:  LINEAR
23     MOLECULE TYPE:  CDNA
24
25 US-08-468-012A-1

```

```

US-09-173-581-10
; Sequence 10, Application US/09173581A
; Patent No. 6013455
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Tang, Y. Tom
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina
; APPLICANT: Azimzai, Yalda
; APPLICANT: Lu, Aina
; TITLE OF INVENTION: Protein Kinase Homologs
; FILE REFERENCE: PF-0614 US
; CURRENT APPLICATION NUMBER: US/09/173,581A
; CURRENT FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 1427
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 119819
US-09-173-581-10

Query Match      5.2%; Score 35.2; DB 3; Length 1427;
Best Local Similarity 53.7%; Pred. No. 0.39;
Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 348 AATGCAGAGAATTTTCATTTTCATGGTGACTTCAATGCTGGTGCAGCTACGTCGCCCAAG 407
DB 857 AATGCAGATAAGATCAAGATCGAGACCTGGGCAACGCTGCTGGGTGCACAAG 916
QY 408 AAGGCTCGAAGGACATCCGCTGAGGACGGACCCCAAGTTCGTTTGGCTGTATCGGGGAC 467
DB 917 CACTTCACGGAAGACATCCAGACTCGGCAGTACCGGGCCGTCGAGGTGCTGATCGGGGCC 976
QY 468 CAAGAGGACACCAACGG 483
DB 977 GAATACGGCCCCCGG 992

RESULT 14
US-09-420-915-10
; Sequence 10, Application US/09420915
; Patent No. 6264947
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Tang, Y. Tom
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina
; APPLICANT: Azimzai, Yalda
; APPLICANT: Lu, Aina
; TITLE OF INVENTION: Protein Kinase Homologs
; FILE REFERENCE: PF-0614 US
; CURRENT APPLICATION NUMBER: US/09/420,915
; CURRENT FILING DATE: 1999-10-20
; EARLIER APPLICATION NUMBER: US 09/173,581
; EARLIER FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 1427
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 119819
US-09-420-915-10

```

```

Query Match      5.2%; Score 35.2; DB 4; Length 1427;
Best Local Similarity 53.7%; Pred. No. 0.39;
Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 348 AATGCAGAGAATTTTCATTTTCATGGTGACTTCAATGCTGGTGCAGCTACGTCGCCCAAG 407
DB 857 AATGCAGATAAGATCAAGATCGAGACCTGGGCAACGCTGCTGGGTGCACAAG 916
QY 408 AAGGCTCGAAGGACATCCGCTGAGGACGGACCCCAAGTTCGTTTGGCTGTATCGGGGAC 467
DB 917 CACTTCACGGAAGACATCCAGACTCGGCAGTACCGGGCCGTCGAGGTGCTGATCGGGGCC 976
QY 468 CAAGAGGACACCAACGG 483
DB 977 GAATACGGCCCCCGG 992

RESULT 15
US-08-924-440-1
; Sequence 1, Application US/08924440
; Patent No. 5871550
; GENERAL INFORMATION:
; APPLICANT: Frits et al.
; TITLE OF INVENTION: MUTANT THERMONOSPORA SPP. CELLULASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International, Inc.
; STREET: 925 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/924,440
; FILING DATE: August 27, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Christopher L. Stone
; REGISTRATION NUMBER: 35,696
; REFERENCE/DOCKET NUMBER: GC388
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 846-7555
; TELEFAX: (650) 845-6504
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1293 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-924-440-1

Query Match      4.9%; Score 32.8; DB 2; Length 1293;
Best Local Similarity 58.0%; Pred. No. 2.1;
Matches 58; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 377 CTTCAATGCTGCTGCTAGCTAGCTCCCAAGAGGCTGGAAGGACATCGGCTGAGGAC 436
DB 1080 CTACACCGGTGACGGCGCAACGACTTCAGATGGCCGCCCTACATCGACCTGATGC 1139
QY 437 GGACCCCAAGTTCGTTTGGCTGATCGGGGACCAAGAGGAC 476
DB 1140 GGAACGGAAGATCGGGTGGACCAAGTGGAACTACTCTCGGAC 1179

Search completed: December 9, 2002, 09:19:58

```

Mon Dec 9 13:45:04 2002

Job time : 57.9082 secs

us-09-905-114-3.rni

---

Page 10

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 09:00:58 ; Search time 133.35 Seconds  
(without alignments)  
1962.092 Million cell updates/sec

Title: US-09-905-114-3  
Perfect score: 671  
Sequence: 1 acaacaggtatgccccata.....atcatcatgaagaaccatga 671

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 350425 seqs, 19496369 residues  
Total number of hits satisfying chosen parameters: 700850

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_NA.\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	671	100.0	671	10	US-09-905-114-3
2	592	88.2	592	10	US-09-905-114-1
3	527	78.5	1023	10	US-09-962-436-336
4	527	78.5	1023	10	US-09-954-456-271
5	527	78.5	1023	10	US-09-880-107-3392
6	527	78.5	1079	12	US-10-133-085-1
7	527	78.5	1079	12	US-10-074-509-1
8	474.8	70.8	1124	12	US-10-133-065-11
9	474.8	70.8	1124	12	US-10-074-509-11
10	195.6	29.2	392	10	US-09-880-107-630
11	163.2	24.3	1039	10	US-09-771-078-1
12	163.2	24.3	1039	10	US-09-825-012-2
13	161.6	24.1	783	10	US-09-825-012-4
14	161.6	24.1	858	10	US-09-825-012-6
15	161.6	24.1	1548	10	US-09-825-012-62
16	161.6	24.1	1548	10	US-09-825-012-63
17	161.6	24.1	1548	10	US-09-825-012-65
18	161.6	24.1	1554	10	US-09-825-012-34
19	161.6	24.1	1554	10	US-09-825-012-35

ALIGNMENTS

RESULT 1  
US-09-905-114-3  
; Sequence 3, Application US/09905114  
; Patent No. US20020048745A1  
; GENERAL INFORMATION:  
; APPLICANT: ZHANG, HUANNIN  
; APPLICANT: AX, ROY L  
; APPLICANT: BELLIN, MARY E  
; TITLE OF INVENTION: ISOLATED POLYNUCLEOTIDE SEQUENCES ENCODING A FERTILITY ASSOCIATED  
; FILE REFERENCE: 21070US20  
; CURRENT APPLICATION NUMBER: US/09/905,114  
; PRIOR FILING DATE: 2001-07-14  
; PRIOR APPLICATION NUMBER: US 60/218,140  
; PRIOR FILING DATE: 2000-07-14  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 671  
; TYPE: DNA  
; ORGANISM: Bos sp.  
US-09-905-114-3

Query Match	100.0%	Score 671;	DB 10;	Length 671;
Best Local Similarity	100.0%	Pred. No. 1.4e-205;		
Mismatches 671;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ACAAAGGATCTGCCCATCTGATGAGAACTAAACGGAATTCAGAAAGGCATAA	60	
Db	1	ACAAAGGATCTGCCCATCTGATGAGAACTAAACGGAATTCAGAAAGGCATAA	60	
Qy	61	CATCAACTATGTGATTAGCTCTCGCTTGGGAGAAACACATATAAAGACAGTATGCCT	120	
Db	61	CATCAACTATGTGATTAGCTCTCGCTTGGGAGAAACACATATAAAGACAGTATGCCT	120	
Qy	121	TTCTCTATAAGAAAGCTAGTGTCTGTAAACAAAGCTACCTCTACACGACTATCAGG	180	
Db	121	TTCTCTATAAGAAAGCTAGTGTCTGTAAACAAAGCTACCTCTACACGACTATCAGG	180	
Qy	181	CTGGAGACCCAGATGTTTTCCAGGAAACCCCTTTGTGCTGTTCCAGTACCCCTACA	240	
Db	181	CTGGAGACCCAGATGTTTTCCAGGAAACCCCTTTGTGCTGTTCCAGTACCCCTACA	240	

QY 241 CCGGTGTAAGACTTCTGATGTTGCCCCCTGCACACACCCCTGAGATCCCTTAGAG 300  
DB 241 CCGGTGTAAGACTTCTGATGTTGCCCCCTGCACACACCCCTGAGATCCCTTAGAG 300  
QY 301 AGATTGATGAGCTGGCTGATGCTCTACACAGATGTGAAACGTGCTGAAATGAGAAATT 360  
DB 301 AGATTGATGAGCTGGCTGATGCTCTACACAGATGTGAAACGTGCTGAAATGAGAAATT 360  
QY 361 TCATTTTCATGAGTGTGACTTCAATGCTGCTGACGTAAGTCCCAAGAAAGCTTGAAGG 420  
DB 361 TCATTTTCATGAGTGTGACTTCAATGCTGCTGACGTAAGTCCCAAGAAAGCTTGAAGG 420  
QY 421 ACATCCGCTGAGAGCGAGACCCCAAGTGGTTGGCTGATGGGGAGCAAGAGACACCA 480  
DB 421 ACATCCGCTGAGAGCGAGACCCCAAGTGGTTGGCTGATGGGGAGCAAGAGACACCA 480  
QY 481 CCGTCAAGAAAGACACAACATGCGCTTATGACAGATCGTGTGAGAGACAAATATTG 540  
DB 481 CCGTCAAGAAAGACACAACATGCGCTTATGACAGATCGTGTGAGAGACAAATATTG 540  
QY 541 TCAACTCTGTGTGCTCTCAATCAAACTGCTTGTGATTTCCAGAAAGCTTACAGTTGT 600  
DB 541 TCAACTCTGTGTGCTCTCAATCAAACTGCTTGTGATTTCCAGAAAGCTTACAGTTGT 600  
QY 601 CTGAATCGAAGGCCCTGATGTCAGCGACCACTTCCAGTTCAATCATCATCATCATG 660  
DB 601 CTGAATCGAAGGCCCTGATGTCAGCGACCACTTCCAGTTCAATCATCATCATCATG 660  
QY 661 AAGAACCATGA 671  
DB 661 AAGAACCATGA 671

RESULT 2  
US-09-905-114-1

; Sequence 1, Application US/09905114  
; Patent No. US20020048745A1  
; GENERAL INFORMATION:  
; APPLICANT: ZHANG, HUAMIN  
; APPLICANT: AX, ROY L  
; APPLICANT: BELLIN, MARY E  
; TITLE OF INVENTION: ISOLATED POLYNUCLEOTIDE SEQUENCES ENCODING A FERTILITY ASSOCIATED  
; FILE REFERENCE: 210707US20  
; CURRENT APPLICATION NUMBER: US/09/905,114  
; PRIOR FILING DATE: 2001-07-14  
; PRIOR APPLICATION NUMBER: US 60/218,140  
; PRIOR FILING DATE: 2000-07-14  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 592  
; TYPE: DNA  
; ORGANISM: Bos sp.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(591)  
; OTHER INFORMATION:  
; US-09-905-114-1

Query Match 88.2%; Score 592; DB 10; Length 592;  
Best Local Similarity 100.0%; Pred. No. 3,1e-180;  
Matches 592; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 GAGAAGCTAAACGGAATTTCAAGAAAGCATATCAATCAACTATGTATTAGCTCTGCG 86  
DB 1 GAGAAGCTAAACGGAATTTCAAGAAAGCATATCAATCAACTATGTATTAGCTCTGCG 60  
QY 87 CTGGAAGAAACATATTAAGAAAGAGATGCTTTCTCTATTAAGAAAGCTAGTGTCT 146  
DB 61 CTGGAAGAAACATATTAAGAAAGAGATGCTTTCTCTATTAAGAAAGCTAGTGTCT 120  
QY 147 GTAAACAAAGCTACCTTACACGACTATCAGGCTGAGAGCGAGATGTGTTTCCAGG 206

DB 121 GTAAACAAAGCTACCTTACACGACTATCAGGCTGAGAGCGAGATGTGTTTCCAGG 180  
QY 207 GAACCCCTTTGTGTGTGTTGCTGATCACTACACCGCTGTCAAGAGACTTGTGATGTC 266  
DB 181 GAACCCCTTTGTGTGTGTTGCTGATCACTACACCGCTGTCAAGAGACTTGTGATGTC 240  
QY 267 CCCCTGCACACACCCTCTGAGACATCCGTTAGAGATTTAGAGCTGCTGATGTTAC 326  
DB 241 CCCCTGCACACACCCTCTGAGACATCCGTTAGAGATTTAGAGCTGCTGATGTTAC 300  
QY 327 ACAGATGAAACCTGCTGATGAGATGACAGAAATTTCAATTTCAAGGTTACTCAATGCT 386  
DB 301 ACAGATGAAACCTGCTGATGAGATGACAGAAATTTCAATTTCAAGGTTACTCAATGCT 360  
QY 387 GCGTGCAGCTACGTTCCCAAGAGGCTTGAAGAGACATCCGCTGAGAGCGAGACCCCAAG 446  
DB 361 GCGTGCAGCTACGTTCCCAAGAGGCTTGAAGAGACATCCGCTGAGAGCGAGACCCCAAG 420  
QY 447 TTGCTTTGGCTGATCGGGGACCAAGAGACACACGCTCAAGAAAGACCAAACTGCGCC 506  
DB 421 TTGCTTTGGCTGATCGGGGACCAAGAGACACACGCTCAAGAAAGACCAAACTGCGCC 480  
QY 507 TATGACAGATCGTCTTGAAGAGCAAAATTTGTCAACTGCTGTTGCTCAATCAAC 566  
DB 481 TATGACAGATCGTCTTGAAGAGCAAAATTTGTCAACTGCTGTTGCTCAATCAAC 540  
QY 567 CTGCTTTGATTTCCAGAAAGCTTACAGTTGTCTGAATGAAAGGCCCTG 618  
DB 541 CTGCTTTGATTTCCAGAAAGCTTACAGTTGTCTGAATGAAAGGCCCTG 592

RESULT 3  
US-09-962-436-336

; Sequence 336, Application US/09962436  
; Patent No. US20020081301A1  
; GENERAL INFORMATION:  
; APPLICANT: Soppet, Daniel  
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu  
; FILE REFERENCE: 689290-75  
; CURRENT APPLICATION NUMBER: US/09/962,436  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,082  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/234,924  
; PRIOR FILING DATE: 2000-09-25  
; NUMBER OF SEQ ID NOS: 568  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 336  
; LENGTH: 1023  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-962-436-336

Query Match 78.5%; Score 527; DB 10; Length 1023;  
Best Local Similarity 87.8%; Pred. No. 3,4e-159;  
Matches 575; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 1 ACAACAGATCTGCCCATACTGATGAGAGCAAGCTTAACGGAATTTCAAGAAAGCATTA 60  
DB 215 ACAACAGATCTGCCCATACTGATGAGAGCAAGCTTAACGGAATTTCAAGAGCATTA 274  
QY 61 CATACAACTATGTATTAGCTCTGCGCTTGAAGAAACATATTAAGAAACAGTATGCT 120  
DB 275 CGTACAACTATGTATTAGCTCTGCGCTTGAAGAAACATATTAAGAAACATATGCT 334  
QY 121 TTCTCTTAAGAAAGCACTAGTGTCTGTAAGAAACAAACTACTCTTACACGACTATCAG 180  
DB 335 TTCTCTTAAGAAAGCACTAGTGTCTGTAAGAGAGTTTCACTACCATGACTATCAG 394  
QY 181 CTGAGAGCGAGATGTGTTTCCAGGAGACCTTTGAGTCTGTTGCTCAAGTACCCCTCA 240  
DB 395 ATGAGAGCGAGATGTGTTTCCAGGAGACCTTTGAGTCTGTTGCTCAATCTCCCA 454



Best Local Similarity 87.8%; Pred. No. 3.4e-159;  
Matches 575; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```
QY 1 ACAACAGGATCTGCCCATACTAGATGAGAACTAAACGGAATTTCAAGAAAGGCATTA 60
D 215 ACAACAGGATCTGCCCATACTAGATGAGAACTAAACGGAATTTCAAGAAAGGCATTA 274
QY 61 CATACACTATGTGATTTAGCTCTGCGCTTTGGAAAGAACATATTAAGACGATATGCTT 120
D 275 CGTACAACTATGTGATTTAGCTCTGCGCTTTGGAAAGAACATATTAAGACGATATGCTT 334
QY 121 TTCCTATTAAGAAAGAGTAGTGTCTGTAACCAAGCTACCTTACCAAGCAGTATCAGG 180
D 335 TTCTCTAAGAGAAAGAGTAGTGTCTGTAACCAAGCTACCTTACCAAGCAGTATCAGG 394
QY 181 CTGAGAGCAGATGTGTTTTCAGAGAAACCTTTTGTGCTGTGTTCCAGTACCCCTTACA 240
D 395 ATGGAGACGAGATGTGTTTTCAGAGAAACCTTTTGTGCTGTGTTCCAGTACCCCTTACA 454
QY 241 CCGCTGTCAAGACTTGTGATTTGCTCCCTGCAACCAACCCCTGAGACATTCCTTAAG 300
D 455 CTGCTGTCAAGACTTGTGATTTGCTCCCTGCAACCAACCCCTGAGACATTCCTTAAG 514
QY 301 AGATTGATGAGCTGCTGATGTCTACACAGATGTGAACCGTCTGGAATGACAGAAAT 360
D 515 AGATGATGAGCTGCTGATGTCTACACAGATGTGAACCGTCTGGAATGACAGAAAT 574
QY 361 TCATTTTCATGGGTGACTTCAATGTGCTGAGCTGAGCTGCTCCCAAGAGGCTTGAAG 420
D 575 TCATTTTCATGGGTGACTTCAATGTGCTGAGCTGAGCTGCTCCCAAGAGGCTTGAAG 634
QY 421 ACATCCGCTGAGAGCGAGACCCCAAGTTCGTTGGCTGATCGGGAGCAAGAGGACACCA 480
D 635 ACATCCGCTGAGAGCGAGACCCCAAGTTCGTTGGCTGATCGGGAGCAAGAGGACACCA 694
QY 481 CGGTCAAGAGAGCAAACTACGCTGCTGACAGATCGTGTGAGAGACAAATATTTG 540
D 695 CGGTCAAGAGAGCAAACTACGCTGCTGACAGATCGTGTGAGAGACAAATATTTG 754
QY 541 TCAACTCTGTGTGCTCTCAATCAAACTCGTCTTGTGATTTCCAGAAAGCTTACAGATTGT 600
D 755 TCAACTCTGTGTGCTCTCAATCAAACTCGTCTTGTGATTTCCAGAAAGCTTACAGATTGT 814
QY 601 CTGAATCGAAGGCGCTGATGTCAAGGACCACTTTCAGTTATCATCATCATCA 655
D 815 CTGAATCGAAGGCGCTGATGTCAAGGACCACTTTCAGTTATCATCATCATCA 869
```

RESULT 6  
US-10-133-065-1  
Sequence 1, Application US/10133065  
Patent No. US20020123122A1

## GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
TITLE OF INVENTION: HUMAN DNASE  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/133,065  
FILING DATE: 26-Apr-2002  
CLASSIFICATION: <Unknown>

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/643,520  
FILING DATE: 22/08/2000  
APPLICATION NUMBER: 08/794827  
FILING DATE: 04-Feb-1997  
APPLICATION NUMBER: 60/109796  
FILING DATE: 05-Feb-1996

## ATTORNEY/AGENT INFORMATION:

NAME: Evans, David W  
REGISTRATION NUMBER: \*See attached limited Recognition under 37

## REFERENCE/DOCKET NUMBER:

C.F.R. 10.9(b)  
TELECOMMUNICATION INFORMATION:  
PI000R1C3

## TELEPHONE:

650/225-1739  
TELEFAX: 650/952-9881

## INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1079 base pairs  
TYPE: Nucleic Acid

STRANDEDNESS: Single  
TOPOLOGY: Linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-133-065-1

Query Match 78.5%; Score 527; DB 12; Length 1079;  
Best Local Similarity 87.8%; Pred. No. 3.5e-159;

Matches 575; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```
QY 1 ACAACAGGATCTGCCCATACTAGATGAGAACTAAACGGAATTTCAAGAAAGGCATTA 60
D 261 ACAACAGGATCTGCCCATACTAGATGAGAACTAAACGGAATTTCAAGAAAGGCATTA 320
QY 61 CATACACTATGTGATTTAGCTCTGCGCTTTGGAAAGAACATATTAAGACGATATGCTT 120
D 321 CGTACAACTATGTGATTTAGCTCTGCGCTTTGGAAAGAACATATTAAGACGATATGCTT 380
QY 121 TTCCTATTAAGAAAGAGTAGTGTCTGTAACCAAGCTACCTTACCAAGCAGTATCAGG 180
D 381 TTCCTATTAAGAAAGAGTAGTGTCTGTAACCAAGCTACCTTACCAAGCAGTATCAGG 440
QY 181 CTGAGAGCAGATGTGTTTTCAGAGAAACCTTTGTTGCTGCTGCTCAGTACCCCTTACA 240
D 441 ATGGAGACGAGATGTGTTTTCAGAGAAACCTTTGTTGCTGCTGCTCAGTACCCCTTACA 500
QY 241 CCGCTGTCAAGACTTGTGATTTGCTCCCTGCAACCAACCCCTGAGACATTCCTTAAG 300
D 501 CTGCTGTCAAGACTTGTGATTTGCTCCCTGCAACCAACCCCTGAGACATTCCTTAAG 560
QY 301 AGATTGATGAGCTGCTGATGTCTACACAGATGTGAACCGTCTGGAATGACAGAAAT 360
D 561 AGATGATGAGCTGCTGATGTCTACACAGATGTGAACCGTCTGGAATGACAGAAAT 620
QY 361 TCATTTTCATGGGTGACTTCAATGTGCTGAGCTGAGCTGCTCCCAAGAGGCTTGAAG 420
D 621 TCATTTTCATGGGTGACTTCAATGTGCTGAGCTGAGCTGCTCCCAAGAGGCTTGAAG 680
QY 421 ACATCCGCTGAGAGCGAGACCCCAAGTTCGTTGGCTGATCGGGAGCAAGAGGACACCA 480
D 681 ACATCCGCTGAGAGCGAGACCCCAAGTTCGTTGGCTGATCGGGAGCAAGAGGACACCA 740
QY 481 CCGTCAAGAGAGCAAACTACGCTGCTGACAGATCGTGTGAGAGACAAATATTTG 540
D 741 CCGTCAAGAGAGCAAACTACGCTGCTGACAGATCGTGTGAGAGACAAATATTTG 800
QY 541 TCAACTCTGTGTGCTCTCAATCAAACTCGTCTTGTGATTTCCAGAAAGCTTACAGATTGT 600
D 801 TCAACTCTGTGTGCTCTCAATCAAACTCGTCTTGTGATTTCCAGAAAGCTTACAGATTGT 860
QY 601 CTGAATCGAAGGCGCTGATGTCAAGGACCACTTTCAGTTATCATCATCATCA 655
D 861 CTGAATCGAAGGCGCTGATGTCAAGGACCACTTTCAGTTATCATCATCATCA 915
```





Query Match 70.8% Score 474.8; DB 12; Length 1124;  
 Best Local Similarity 82.9%; Pred. No. 2.1e-142;  
 Matches 542; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

```

    2 CAACGAGATCGCCCCCTACTGATGAGAGGTAACGAAATTCAGAAAAGGCATAC 61
    Db 379 CAACACATCTGTCCCATCTGATGAGAGGTAATGAAATTCACGAGAGACACAC 438
    QY 62 ATACACTATGATGATTAAGCTCTGCTTGAAGAAACATATTAAGAACATATGCTT 121
    Db 439 ATACACTATGATGATTAAGCTCTGCTTGAAGAAACATATTAAGAACATATGCTT 498
    QY 122 TCTGTATTAAGAAAGCTAGTGTCTGTAAACAAAGCTACCTTACCGAGATACAGC 181
    Db 499 GGTCAAGAGAGAGAGCTGTGTCTGTAAACAAAGCTACCTTACCGAGATACAGC 558
    QY 182 TGGAGACGAGATGATGATTTCCAGGAAACCTTTGTGCTGATTCAGTCACTTAC 241
    Db 559 TGGAGACACAGACGATTTTCCAGGAGACCTTTGTGCTGATTCAGTCACTTAC 618
    QY 242 CGCTGTCAAGAGCTTGTGATTTGCTCCCTGACACACCCCTGAGACATCCGTTAGA 301
    Db 619 TGCTGTCAAGAGCTTGTGATTTGCTCCCTGACACACCCCTGAGACATCCGTTAGA 678
    QY 302 GATGATAGCTGCTGATGTTTACACAGATGTGAACGCTGCTGGAATGACAGAAATTT 361
    Db 679 GATGATAGCTGCTGATGTTTACACAGATGTGAACGCTGCTGGAATGACAGAAATTT 738
    QY 362 CATTTTCATGAGGCTGCTGATGTTTACAGTACGCTGCTGCTGCTGCTGCTGCTGCTG 421
    Db 739 CATTTTCATGAGGCTGCTGATGTTTACAGTACGCTGCTGCTGCTGCTGCTGCTGCTG 798
    QY 422 CATCGCTGAGAGACGACCCCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 481
    Db 799 CATCGCTGAGAGACGACCCCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 858
    QY 482 GGTCAAGAGAGGACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 541
    Db 859 GGTCAAGAGAGGACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 918
    QY 542 CAACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 601
    Db 919 CAACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 978
    QY 602 TGAATGAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 655
    Db 979 TGAATGAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1032
    
```

RESULT 9  
 US-10-074-509-11  
 Sequence 11, Application US/10074509  
 Patent No. US20020142437A1  
 GENERAL INFORMATION:  
 APPLICANT: Baker, Kevin P.  
 Attorney: Baton, Will F.  
 TITLE OF INVENTION: HUMAN DNASE  
 NUMBER OF SEQUENCES: 11  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: Genentech, Inc.  
 STREET: 1 DNA Way  
 CITY: South San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94080  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: WinPatIn (Genentech)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/074,509  
 FILING DATE: 11-Feb-2002

US-10-074-509-11  
 Sequence 11, Application US/10074509  
 Patent No. US20020142437A1  
 GENERAL INFORMATION:  
 APPLICANT: Baker, Kevin P.  
 Attorney: Baton, Will F.  
 TITLE OF INVENTION: HUMAN DNASE  
 NUMBER OF SEQUENCES: 11  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: Genentech, Inc.  
 STREET: 1 DNA Way  
 CITY: South San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94080  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: WinPatIn (Genentech)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/074,509  
 FILING DATE: 11-Feb-2002

CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 09/643,520  
 FILING DATE: 22/08/2000  
 APPLICATION NUMBER: 08/794827  
 FILING DATE: 04-Feb-1997  
 APPLICATION NUMBER: 60/109796  
 FILING DATE: 05-Feb-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Johnston, Sean A.  
 REGISTRATION NUMBER: 35,910  
 REFERENCE/DOCKET NUMBER: P100OR1C2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 650/225-3562  
 TELEFAX: 650/952-9881  
 INFORMATION FOR SEQ ID NO: 11:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1124 base pairs  
 TYPE: Nucleic Acid  
 STRANDEDNESS: Single  
 TOPOLOGY: Linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 11:  
 US-10-074-509-11

Query Match 70.8% Score 474.8; DB 12; Length 1124;  
 Best Local Similarity 82.9%; Pred. No. 2.1e-142;  
 Matches 542; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

```

    2 CAACGAGATCGCCCCCTACTGATGAGAGGTAACGAAATTCAGAAAAGGCATAC 61
    Db 379 CAACACATCTGTCCCATCTGATGAGAGGTAATGAAATTCACGAGAGACACAC 438
    QY 62 ATACACTATGATGATTAAGCTCTGCTTGAAGAAACATATTAAGAACATATGCTT 121
    Db 439 ATACACTATGATGATTAAGCTCTGCTTGAAGAAACATATTAAGAACATATGCTT 498
    QY 122 TCTGTATTAAGAAAGCTAGTGTCTGTAAACAAAGCTACCTTACCGAGATACAGC 181
    Db 499 GGTCAAGAGAGAGCTGTGTCTGTAAACAAAGCTACCTTACCGAGATACAGC 558
    QY 182 TGGAGACGAGATGATGATTTCCAGGAAACCTTTGTGCTGATTCAGTCACTTAC 241
    Db 559 TGGAGACACAGACGATTTTCCAGGAGACCTTTGTGCTGATTCAGTCACTTAC 618
    QY 242 CGCTGTCAAGAGCTTGTGATTTGCTCCCTGACACACCCCTGAGACATCCGTTAGA 301
    Db 619 TGCTGTCAAGAGCTTGTGATTTGCTCCCTGACACACCCCTGAGACATCCGTTAGA 678
    QY 302 GATGATAGCTGCTGATGTTTACACAGATGTGAACGCTGCTGGAATGACAGAAATTT 361
    Db 679 GATGATAGCTGCTGATGTTTACACAGATGTGAACGCTGCTGGAATGACAGAAATTT 738
    QY 362 CATTTTCATGAGGCTGCTGATGTTTACAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 421
    Db 739 CATTTTCATGAGGCTGCTGATGTTTACAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 798
    QY 422 CATCGCTGAGAGACGACCCCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 481
    Db 799 CATCGCTGAGAGACGACCCCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 858
    QY 482 GGTCAAGAGAGGACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 541
    Db 859 GGTCAAGAGAGGACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 918
    QY 542 CAACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 601
    Db 919 CAACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 978
    QY 602 TGAATGAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 655
    Db 979 TGAATGAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1032
    
```

## RESULT 10

US-09-880-107-630/c  
; Sequence 630, Application US/09880107  
; Patent No. US20020142981A1  
; GENERAL INFORMATION:  
; APPLICANT: Horne, Darci T.  
; APPLICANT: Vockrey, Joseph G.  
; APPLICANT: Scherf, Uwe  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
; CURRENT APPLICATION NUMBER: US/09/880,107  
; CURRENT FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/211,379  
; PRIOR FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: US 60/237,054  
; PRIOR FILING DATE: 2000-10-02  
; NUMBER OF SEQ ID NOS: 3950  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 630  
; LENGTH: 392  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA279676  
US-09-880-107-630

Query Match 29.2%; Score 195.6; DB 10; Length 392;  
Best Local Similarity: 84.9%; Pred. No. 5.4e-53;  
Matches 219; Conservative 0; Mismatches 39; Indels 0; Gaps 0;  
QY 398 CGTCCCAAGAGCGCTGGAAGACATCCGCTTGAGGACGACCCCAAGTTCTGTTGGCT 457  
Db 392 CGTCCCAAGAGCGCTGGAAGACATCCGCTTGAGGACTGACCCCAAGTTCTGTTGGCT 333  
QY 458 GATCGGGGACCAAGAGACACACCGTCAAGAGAGACCAAACTGCGCTATGACAGGAT 517  
Db 332 GATCGGGGACCAAGAGAGACACACCGTCAAGAGAGACCAAACTGTCATATGACAGGAT 273  
QY 518 CGTGCTTAGAGGACAAATATCTCAACTCTGGTGTCTCTCAATCAAACTCGTCTTTGA 577  
Db 272 TGTGCTTAGAGGACAAATATCTCAACTCTGGTGTCTCTCAATCAAACTCGTCTTTGA 213  
QY 578 TTTCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCGCTGGATGTGAGGACCACTTTCC 637  
Db 212 TTTCAGAAAGCTTACAGGTTGTCTGAATCGAAGGCGCTGGATGTGAGGACCACTTTCC 153  
QY 638 AGTTCAATCATCATCA 655  
Db 152 AGTTGAATTTAACTACA 135

## RESULT 11

US-09-771-078-1  
; Sequence 1, Application US/09771078  
; Patent No. US20020034727A1  
; GENERAL INFORMATION:  
; APPLICANT: Mtsny, Randall J.,  
; APPLICANT: Daugherty, Ann L.,  
; APPLICANT: Patapoff, Thomas W.  
; TITLE OF INVENTION: DNase Compaction Assay  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/771,078  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/971,019  
ATTORNEY/AGENT INFORMATION:  
NAME: Glaister, Debra J.  
REGISTRATION NUMBER: 33,888  
REFERENCE/DOCKET NUMBER: 792  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-3177  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1039 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-771-078-1

Query Match 24.3%; Score 163.2; DB 10; Length 1039;  
Best Local Similarity 57.4%; Pred. No. 2.7e-42;  
Matches 335; Conservative 0; Mismatches 243; Indels 6; Gaps 2;  
QY 60 ACATACAATATGTGATTAGCTCTCGCTTGGAGAAACACATATATAAGAACAGTATGCC 119  
Db 409 ACCTATCATCTAGCTGTGCTAGTGAGCCACTGGGACGAAACAGCTATATAAGGAGCCTACCTG 468  
QY 120 TTTCTCTATAAAGAAAGCTAGTGTCTGTAAAAAAGCTACCTCTACCAACGA---CTAT 176  
Db 469 TTCTGTACAGGCTGACAGGTGTCTGCGGTGGACAGTACTACTACGATGTGCTGC 528  
QY 177 CAGGCTGGAGACGAGATGTTTTCCAGGAAACCTTTTGTGTCTGTTTCCAGTCACTCC 236  
Db 529 GAGCCCTGCGGAAACGACACCTTCAACCGAGAGCCAGCCATTGTCTCAGGTTCTTCTCCCGG 588  
QY 237 TACACCGCTGTCAAGACTTCGTGATTGTCCCTTCACACACCCCTGAGACATCGGTT 296  
Db 589 TTACAGAGGTGAGGAGTTTCCCATTTGTCCCTGCTATGCGGCGCCCGGGGACGAGTA 648  
QY 297 AGAGAGATGATGAGCTGCTGATGTCTACACAGATGTGAAACGTCTGGAATGCAGAG 356  
Db 649 GCGGAGATCGAGCTCTCTATGACGTCTACCTGGATGTCCAGAGAAATGGGCTTGGAG 708  
QY 357 AATTTCAATTTTCATGGGTGACTTCAATGTGCTGCTGAGTACGTCTCCCAAGAGCGCTGG 416  
Db 709 GAGCTCATGTTGATGGCGACTTCAATGCGGCTGCAGCTATGTGAGACCTCTCCAGTGG 768  
QY 417 AAGGACATCGGCTGAGGACGACCCCAAGTTGTTGGCTGATCGGGGACCAAGAGGAC 476  
Db 769 TCATCCATCCGCTGTGGACAAGCCCACTTTCCAGTGGCTGATCTCCCGACGAGCGCTGAC 828  
QY 477 ACCACGTTCAAGAAGAGACAAACTGCGCTATGACAGGATGCTGCTTAGAGAGCAAAAT 536  
Db 829 ACACAG---CTACACCCACGCTGCTGCTATGACAGGATCTGTTGTCAGGATGCTG 885  
QY 537 ATTTGCAACTCTGGTGGTCTCAATCAAAACCTCGTCTTTGATTTCCAGAAAGCTTACAGG 596  
Db 886 CTCGAGGCGGCTTTGTTCCCGACTCGGCTCTTCCCTTTAACTTCCAGGCTGCTTATGGC 945  
QY 597 TTGCTGTAATCGAAGCCCTGGATGTACGAGCACTTTTCCAGT 640  
Db 946 CTGAGTGACCAACTGGCCCCAAGCCCATCAGTGACCACTATCCAGT 989

## RESULT 12

US-09-825-012-2  
; Sequence 2, Application US/09825012  
; Patent No. US20020122798A1

```

; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1039
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-825-012-2

Query Match      24.3%; Score 163.2; DB 10; Length 1039;
Best Local Similarity 57.4%; Pred. No. 2.7e-42;
Matches 335; Conservative 0; Mismatches 243; Indels 6; Gaps 2;

QY 60 ACATACAACATATGATGATGCTCTGCTTGGAGAAACACATATATAAGAAACAGTATGCC 119
Db 409 ACCATATCACTAGTGTGATGAGCACTGGAGCGAAACAGCTATAGAGAGCGTACCTG 468
QY 120 TTTCTCTATAAAGAAAGCTAGTGTCTGTAACAAAGCTACCTTACACGA--CTAT 176
Db 469 TTCGTGTACAGGCTGACAGGTGTCTGCGTGACAGCTACTACTACAGATGATGCTGC 528
QY 177 CAGGCTGAGAGCGAGATGTGTTTCCAGGGAAACCTTTGGGTGTGTTCCAGTACACC 236
Db 529 GAGCCTCGCGGAGACGACCTTCAACGAGAGCCAGCATTTGATGATTTCTTCCCGG 588
QY 237 TACACCGCTGTCAAGACCTTGTGATGTGCTCCCTGCACACACCCCTGAGACATCCGT 296
Db 589 TTACAGAGGTGACAGGAGTTTGCCATTGTCCCTGCATGAGGCGCCCGGGAGCGCAGTA 648
QY 297 AGAGAGATTGATGAGCTGGCTGATCTTACACAGATGTGAAACGTGCTGGAATGCAGAG 356
Db 649 GCCGAGATCGACGCTCTATGACGTCTACCTGATGTCCAAAGAAATGGGGCTTGGAG 708
QY 357 AATTTCAATTTTCAATGGGTGATTCATATGCTGCTGACGTAAGTCCCAAGAAGCGCTGG 416
Db 709 GACGTCAATGTTGATGAGGCGCATTCATATCGGAGTGCAGCTATGTGAGACCTTCCAGTGG 768
QY 417 AAGACATCCGCTGAGAGACGAGACCCCAAGTTGCTTTGGCTGATCGGGACCAAGAGGAC 476
Db 769 TCATTCATCCGCTGTGTGACAAAGCCCACTTCCAGTGGTGTATCCCGACAGGCGCTGAC 828
QY 477 ACCACGCTCAAGAAAGACAAACTGCGCCTTATGACAGGATCGTCTTAAAGACAAAT 536
Db 829 ACCACAG--CTACACCCAGCAGCTGTGCTATGACAGGATCGTGTGAGAGAGTGTG 885
QY 537 ATTGTCAACTGTGTGCTGCTCAATCAAACTGCTTTGATTTCCAGAAAGCTTACAGG 596
Db 886 CTCGAGGCGCGCTTTGTTCCGACTCGGCTTTCCTTTAACTTCCAGGCTGCTATGGC 945
QY 597 TTGTCTGAATCGAAGCGCTGATGTCAGCAGCACTTTCAGT 640
Db 946 CTGAGTGACCAACTGGCCCAAGCCATCAGTACCACTATCCAGT 989

RESULT 13
US-09-825-012-4
; Sequence 4, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03

```

```

; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 783
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-825-012-4

Query Match      24.1%; Score 161.6; DB 10; Length 783;
Best Local Similarity 57.2%; Pred. No. 7.2e-42;
Matches 334; Conservative 0; Mismatches 244; Indels 6; Gaps 2;

QY 60 ACATACAACATATGATGATGCTCTGCTTGGAGAAACACATATATAAGAAACAGTATGCC 119
Db 184 ACCATATCACTAGTGTGATGAGCACTGGAGCGAAACAGCTATAGAGAGCGTACCTG 243
QY 120 TTTCTCTATAAAGAAAGCTAGTGTCTGTAACAAAGCTACCTTACACGA--CTAT 176
Db 244 TTCGTGTACAGGCTGACAGGTGTCTGCGTGAGAGCTACTACTACAGATGATGCTGC 303
QY 177 CAGGCTGAGAGCGAGATGTGTTTCCAGGGAACCTTTGTGCTGTGTTCACTACACC 236
Db 304 GAGCCTCGCGGAGACGACCTTCAACGAGAGCCAGCATTTGTCAGGTTCTTCCCGG 363
QY 237 TACACCGCTGTCAAGACCTTGTGATGTGCTCCCTGCACACACCCCTGAGACATCCGTT 296
Db 364 TTACAGAGGTGACAGGAGTTTGCCATTGTTCCTTGAATGAGGCTCCGGGGAGCGCAGTA 423
QY 297 AGAGAGATTGATGAGCTGGCTGATGTCTACACAGATGTGAAACGCTGGAATGCAGAG 356
Db 424 GCCGAGATCGACGCTCTATGAGTGTACCTGATGTCCAAAGAAATGGGCTTGGAG 483
QY 357 AATTTCAATTTTCAATGGGTGATTCATATGCTGCTGACGTAAGTCCCAAGAAGCGCTGG 416
Db 484 GACGTCAATGTTGATGAGGCGCTTCAATGCGGCTGACGCTATGAGAACCTCCAGTGG 543
QY 417 AAGACATCCGCTGAGAGACGAGACCCCAAGTTGCTTTGGGTGATCGGGACCAAGAGGAC 476
Db 544 TCATTCATCCGCTGTGTGACAAAGCCCACTTCCAGTGGTGTATCCCGACAGGCGTGGAC 603
QY 477 ACCACGCTCAAGAAAGACCAAACTGCGCCTTATGACAGATGTGCTTAAAGACCAAAAT 536
Db 604 ACCACAG--CTACACCCAGCAGCTGTGCTTATGACAGGATCGTGTGAGAGTGTG 660
QY 537 ATTGTCAACTGTGTGCTGCTCAATCAAACTGCTTTGATTTCCAGAAAGCTTACAGG 596
Db 661 CTCGAGGCGCGCTTTGTTCCGACTCGGCTTTCCTTTAACTTCCAGGCTGCTATGGC 720
QY 597 TTGTCTGAATCGAAGCGCTGATGTCAGCAGCACTTTCAGT 640
Db 721 CTGAGTGACCAACTGGCCCAAGCCATCAGTACCACTATCCAGT 764

RESULT 14
US-09-825-012-6
; Sequence 6, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1

```

SEQ ID NO 6  
LENGTH: 858  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-825-012-6

Query Match 24.1%; Score 161.6; DB 10; Length 858;  
Best Local Similarity 57.2%; Pred. No. 7.6e-42;  
Matches 334; Conservative 0; Mismatches 244; Indels 6; Gaps 2;

QY 60 ACATACAACTATGTGATTAGCTCTCGCCTTGGAGAAACACATATAAAGAACAGTATGCC 119  
DB 259 ACCTATCACTAGCTGTGCTAGTGGACCACTGGAGCGAAGAGCTATAGAGGCGCTACCTG 318  
QY 120 TTTCTCTATAAAGAAAGCTAGTGTCTGTAACAAAGCTACTCTACACGAA---CTAT 176  
DB 319 TTCGTGTACAGGCGCTGACAGGTGTCTGGGTGGAGCAGCTACTACTACGATGATGCTGC 378  
QY 177 CAGGCTGGAGAGCGAGATGTGTTTCCAGGGAACCTTTGTGGTCTGGTTCAGTCACCC 236  
DB 379 GAGCCCTGCGGGAACGACACCTTCAACCGAGAGCCAGCCATTGTTCAGGTTCTTCTCCCG 438  
QY 237 TACACCGCTGTCAAGGACTTGTGATTGTCCCTGCACACACCCCTGAGACATCCGTT 296  
DB 439 TTCACAGAGGTGAGGAGTTTGCATTGTTCCCTGCATGCGGCCCCGGGGAGCGCAGTA 498  
QY 297 AGAGAGATTGATGAGCTGGCTGTATGCTACACAGATGTGAACGCTGCTGGAATGCAGAG 356  
DB 499 GCCGAGATCGACGCTCTCTATGACGCTTACCTGGATGTCAGAGAAATGGGGCTTGGAG 558  
QY 357 AATTTCATTTTCACTGGTGACTTCAATGTGGCTGAGCTAGCTTCCCAAGAGGCGCTGG 416  
DB 559 GACGTCATGTTGATGGGCACTTCAATGCGGGCTGCGAGCTATGTGAGACCCCTCCAGTGG 618  
QY 417 AAGGACATCCGCTGAGGAGCGACCCCAAGTTCGTTGGCTGATCGGGACCAAGAGGAC 476  
DB 619 TCATCCATCCGCTGTGGACAAGCCCCACCTTCAGTGGCTGATCCCCAGACGCTGAC 678  
QY 477 ACCAGGTCAGAGAGAGCAAACTGCGCTATGACAGGATCGTGTAGAGGACAAAAT 536  
DB 679 ACCACAG---CTACACCCAGCACTGTGCTATGACAGGATCGTGTGACGGATGCTG 735  
QY 537 ATTGTCAACTCTGGTGGCTCAATCAACCTCGCTTTTGTATTCAGAAAGCTTACAGG 596  
DB 736 CTCGAGGGGCGGTTGTTCCGACCTGGCTCTTCCCTTTAACTTCAGGCTGCCTATGCG 795  
QY 597 TTGTCCTGAATCGAAGCGCTGATGTACGACCACTTTCAGT 640  
DB 796 CTGAGTGACCAACTGGCCCAAGCCCATCAGTGACCACTATCCAGT 839

## RESULT 15

US-09-825-012-62  
Sequence 62, Application US/09825012  
Patent No. US20020122798A1  
GENERAL INFORMATION:  
APPLICANT: Young, Robert  
TITLE OF INVENTION: Compounds for Targeting  
FILE REFERENCE: 43191-256808  
CURRENT APPLICATION NUMBER: US/09/825,012  
CURRENT FILING DATE: 2001-04-03  
PRIOR APPLICATION NUMBER: US 60/237,159  
PRIOR FILING DATE: 2000-10-02  
PRIOR APPLICATION NUMBER: GB 0008049.9  
PRIOR FILING DATE: 2000-04-03  
NUMBER OF SEQ ID NOS: 102  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 62  
LENGTH: 1548  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Humanised HMFG1 Fd - Dnaase I fusion

US-09-825-012-62

Query Match 24.1%; Score 161.6; DB 10; Length 1548;  
Best Local Similarity 57.2%; Pred. No. 1.1e-41;  
Matches 334; Conservative 0; Mismatches 244; Indels 6; Gaps 2;

QY 60 ACATACAACTATGTGATTAGCTCTCGCCTTGGAGAAACACATATAAAGAACAGTATGCC 119  
DB 949 ACCTATCACTAGCTGTGCTAGTGGACCACTGGAGCGAAGAGCTATAGAGGCGCTACCTG 1008  
QY 120 TTTCTCTATAAAGAAAGCTAGTGTCTGTAACAAAGCTACTCTACACGAA---CTAT 176  
DB 1009 TTCGTGTACAGGCGCTGACAGGTGTCTGGGTGGAGCAGCTACTACTACGATGATGCTGC 1068  
QY 177 CAGGCTGGAGAGCGAGATGTGTTTCCAGGGAACCCCTTTGTGGTCTGGTTCAGTCACCC 236  
DB 1069 GAGCCCTGCGGGAACGACACCTTCAACCGAGAGCCAGCCATTGTTCAGGTTCTTCTCCCG 1128  
QY 237 TACACCGCTGTCAAGGACTTGTGATTGTCCCTGCACACACCCCTGAGACATCCGTT 296  
DB 1129 TTCACAGAGGTGAGGAGTTTGCATTGTTCCCTGCATGCGGCCCCGGGGAGCGCAGTA 1188  
QY 297 AGAGAGATTGATGAGCTGGCTGTATGCTACACAGATGTGAACGCTGCTGGAATGCAGAG 356  
DB 1189 GCCGAGATCGACGCTCTCTATGACGCTTACCTGGATGTCAGAGAAATGGGGCTTGGAG 1248  
QY 357 AATTTCATTTTCACTGGTGACTTCAATGTGGCTGCGAGCTACGTCCTCCCAAGAGGCGCTGG 416  
DB 1249 GACGTCATGTTGATGGGCACTTCAATGGGGCTGCGAGCTATGTGAGACCCCTCCAGTGG 1308  
QY 417 AAGGACATCCGCTGAGGAGCGACCCCAAGTTCGTTGGCTGATCGGGACCAAGAGGAC 476  
DB 1309 TCATCCATCCGCTGTGGACAAGCCCCACCTTCCAGTGGCTGATCCCCAGACGCTGAC 1368  
QY 477 ACCAGGTCAGAGAGAGCAAACTGCGCTATGACAGGATCGTGTAGAGGACAAAAT 536  
DB 1369 ACCACAG---CTACACCCAGCACTGTGCTTATGACAGGATCGTGTGACGGGATGCTG 1425  
QY 537 ATTGTCAACTCTGGTGGTCCCTCAATCAAACTCGCTTTTGTATTCAGAAAGCTTACAGG 596  
DB 1426 CTCGAGGGGCGGTTGTTCCGACCTCGGCTCTTCCCTTTAACTTCAGGCTGCCTATGCG 1485  
QY 597 TTGTCCTGAATCGAAGGCGCTGATGTACAGGACCACTTTCAGT 640  
DB 1486 CTGAGTGACCAACTGGCCCAAGCCCATCAGTGACCACTATCCAGT 1529

Search completed: December 9, 2002, 09:24:18  
Job time : 144.35 secs

**This Page Blank (uspto)**